53525

Delaval, Jan

From: DeCloux, Amy
Sent: Tuesday, October 23, 2001 1:20 PM
To: Delaval, Jan
Subject: 09/625,963

Hi cheerful Jan:

Sorry I forgot to send you this search request with the big batch:

Could you please do an open and closed search of the small amino acid sequences of SEQ ID NO:1, SEQ ID NO:2 and SEQ ID NO:3????

Thank you Jan!

Amy

Amy DeCloux
Art Unit 1644
CM1 9B09 Office
CM1 9E12 Mail
(703) 306-5821

Point of Contact:
Jan Delaval
Librarian-Physical Sciences
CM1 1E01 Tel: 308-4498

U.S. DEPARTMENT OF COMMERCE
Patent and Trademark Office

SEARCH REQUEST FORM

53525

Requestor's Name: ____________ Serial Number: ____________

Date: ____________ Phone: ____________ Art Unit: ____________

Search Topic:

Please write a detailed statement of search topic. Describe specifically as possible the subject matter to be searched. Define any terms that may have a special meaning. Give examples or relevant citations, authors keywords, etc., if known. For sequences, please attach a copy of the sequence. You may include a copy of the broadest and/or most relevant claim(s).

BEST AVAILABLE C

Point of Contact:
Jan Delaval
Librarian-Physical Sciences
CM1 1E01 Tel: 308-4498

STAFF USE ONLY

Date completed: [illegible]
Searcher: ____________
Terminal time: ____________
Elapsed time: ____________
CPU time: ____________
Total time: ____________
Number of Searches: ____________
Number of Databases: ____________

Search Site
- ____ STIC
- ____ CM-1
- ____ Pre-S

Type of Search
- ____ N.A. Sequence
- [✓ A.A. Sequence]
- ____ Structure
- ____ Bibliographic

Vendors
- ____ IG Suite
- ____ STN
- ____ Dialog
- ____ APS
- ____ Geninfo
- ____ SDC
- ____ DARC/Questel
- [✓ Other]

```
A;Accession: S33559
A;Molecule type: mRNA
A;Residues: 1-854 <LIW2>
A;Cross-references: EMBL:Z14146; NID:g6903; PID:g6904
C;Genetics:
A;Introns: 14/2; 68/1; 129/3; 151/3; 182/1; 255/3; 361/2; 660/1; 790/3
C;Keywords: alternative initiators
F;1-854/Product: unc-33 protein (long form) #status predicted <MAT1>
F;176-854/Product: unc-33 protein (intermediate form) #status predicted <MAT2>
F;332-854/Product: unc-33 protein (short form) #status predicted <MAT3>


  Query Match             72.5%;  Score 37;  DB 2;  Length 854;
  Best Local Similarity   66.7%;  Pred. No. 41;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        1 RMFPNAPYL 9
            |: ||:|||
Db      773 RLSPNSPYL 781


RESULT  11
A55478
neuronal apoptosis inhibitory protein - human
N;Alternate names: NAIP
C;Species: Homo sapiens (man)
C;Date: 05-Jan-1996 #sequence_revision 05-Jan-1996 #text_change 02-Feb-2001
C;Accession: A55478
R;Roy, N.; Mahadevan, M.S.; McLean, M.; Shutler, G.; Yaraghi, Z.; Farahani, R.; Baird, S
d, T.O.; de Jong, P.J.; Surh, L.; Ikeda, J.E.; Korneluk, R.G.; MacKenzie, A.
Cell 80, 167-178, 1995
A;Title: The gene for neuronal apoptosis inhibitory protein is partially deleted in indi
A;Reference number: A55478; MUID:95112344
A;Accession: A55478
A;Molecule type: mRNA
A;Residues: 1-1232 <ROY>
A;Cross-references: GB:U19251
C;Genetics:
A;Gene: GDB:SMA@; SMA
A;Cross-references: GDB:120378; OMIM:600354; OMIM:253300
A;Map position: 5q12.2-5q13
C;Keywords: apoptosis; ATP; glycoprotein; nucleotide binding; P-loop; transmembrane prot
F;94-110/Domain: transmembrane #status predicted <TMM1>
F;470-477/Region: nucleotide-binding motif A (P-loop)
F;479-496/Domain: transmembrane #status predicted <TMM2>
F;476/Binding site: ATP (Lys) #status predicted
F;618,632,823,923,1035/Binding site: carbohydrate (Asn) (covalent) #status predicted


  Query Match             72.5%;  Score 37;  DB 2;  Length 1232;
  Best Local Similarity   66.7%;  Pred. No. 61;
  Matches    6;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy        1 RMFPNAPYL 9
            | |||.|:|
Db      337 RCFPNCPFL 345


RESULT  12
S17463
glutathione transferase (EC 2.5.1.18) class mu chain Yb5 - mouse (fragments)
C;Species: Mus musculus (house mouse)
C;Date: 22-Nov-1993 #sequence_revision 08-Nov-1996 #text_change 16-Feb-1997
C;Accession: S17463
R;Hayes, J.D.; Kerr, L.A.; Peacock, S.D.; Cronshaw, A.D.; McLellan, L.I.
Biochem. J. 277, 501-512, 1991
A;Title: Hepatic glutathione S-transferases in mice fed on a diet containing the anticar
adient elution of the glutathione-Sepharose affinity matrix.
A;Reference number: S16933; MUID:91315425
A;Accession: S17463
A;Molecule type: protein
A;Residues: 1-38;39-64;65-85 <HAY>
C;Superfamily: glutathione transferase
C;Keywords: blocked amino end; dimer; transferase


  Query Match             70.6%;  Score 36;  DB 2;  Length 85;
  Best Local Similarity   85.7%;  Pred. No. 4.9;
  Matches    6;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        3 FPNAPYL 9
            ||| |||
Db       23 FPNLPYL 29


RESULT  13
S39393
glutathione transferase (EC 2.5.1.18) 3 - rat (fragments)
C;Species: Rattus norvegicus (Norway rat)
C;Date: 20-May-1994 #sequence_revision 10-Nov-1995 #text_change 07-May-1999
C;Accession: S39393
R;Liu, L.F.; Hong, J.L.; Tsai, S.P.; Hsieh, J.C.; Tam, M.F.
Biochem. J. 296, 189-197, 1993
A;Title: Reversible modification of rat liver glutathione S-transferase 3-3 with 1-ch
A;Reference number: S39393; MUID:94071824
A;Accession: S39393
A;Molecule type: protein
A;Residues: 1-115 <LIU>
C;Superfamily: glutathione transferase
C;Keywords: transferase


  Query Match             70.6%;  Score 36;  DB 2;  Length 115;
  Best Local Similarity   85.7%;  Pred. No. 6.9;
  Matches    6;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        3 FPNAPYL 9
            ||| |||
Db       31 FPNLPYL 37


RESULT  14
S17462
glutathione transferase (EC 2.5.1.18) class mu chain Yb2 - mouse (fragments)
C;Species: Mus musculus (house mouse)
C;Date: 22-Nov-1993 #sequence_revision 08-Nov-1996 #text_change 16-Feb-1997
C;Accession: S17462
R;Hayes, J.D.; Kerr, L.A.; Peacock, S.D.; Cronshaw, A.D.; McLellan, L.I.
Biochem. J. 277, 501-512, 1991
A;Title: Hepatic glutathione S-transferases in mice fed on a diet containing the anti
adient elution of the glutathione-Sepharose affinity matrix.
A;Reference number: S16933; MUID:91315425
A;Accession: S17462
A;Molecule type: protein
A;Residues: 1-65;66-128;129-135;136-142 <HAY>
C;Superfamily: glutathione transferase
C;Keywords: dimer; transferase


  Query Match             70.6%;  Score 36;  DB 2;  Length 142;
  Best Local Similarity   85.7%;  Pred. No. 8.7;
  Matches    6;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        3 FPNAPYL 9
            ||| |||
Db       56 FPNLPYL 62


RESULT  15
JX0095
glutathione transferase (EC 2.5.1.18) b - guinea pig
N;Alternate names: glutathione S-alkyltransferase; glutathione S-aryltransferase; glu
C;Species: Cavia porcellus (guinea pig)
C;Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 18-Jun-1993
C;Accession: JX0095
R;Kamei, K.; Oshino, R.; Hara, S.
```

```
  Best Local Similarity   75.0%;  Pred. No. 2.9;
  Matches    6;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RMFPNAPY 8
           |:|||:||
Db     151 RVFPNSPY 158


RESULT   6
C71375
conserved hypothetical integral membrane protein TP0033 - syphilis spirochete
C;Species: Treponema pallidum subsp. pallidum (syphilis spirochete)
C;Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 05-Nov-1999
C;Accession: C71375
R;Fraser, C.M.; Norris, S.J.; Weinstock, G.M.; White, O.; Sutton, G.G.; Dodson, R.; Gwin
rson, J.; Khalak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Utterback, T.; McDo
they, L.; Weidman, J.; Smith, H.O.; Venter, J.C.
Science 281, 375-388, 1998
A;Title: Complete genome sequence of Treponema pallidum, the syphilis spirochete.
A;Reference number: A71250; MUID:98332770
A;Accession: C71375
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-203 <COL>
A;Cross-references: GB:AE001188; GB:AE000520; NID:g3322282; PIDN:AAC65028.1; PID:g332228
A;Experimental source: strain Nichols
C;Genetics:
A;Gene: TP0033


  Query Match             76.5%;  Score 39;  DB 2;  Length 203;
  Best Local Similarity   85.7%;  Pred. No. 3.4;
  Matches    6;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3 FPNAPYL 9
             ||||||:
Db      63 FPNAPYI 69


RESULT   7
F83495
probable hydrolase PA1202 [imported] - Pseudomonas aeruginosa (strain PAO1)
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C;Accession: F83495
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrener, P.; Hickey, M.J.; Br
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathc
A;Reference number: A82950; MUID:20437337
A;Accession: F83495
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-205 <STO>
A;Cross-references: GB:AE004550; GB:AE004091; NID:g9947122; PIDN:AAG04591.1; GSPDB:GN001
A;Experimental source: strain PAO1
C;Genetics:
A;Gene: PA1202


  Query Match             74.5%;  Score 38;  DB 2;  Length 205;
  Best Local Similarity   75.0%;  Pred. No. 5.4;
  Matches    6;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 MFPNAPYL 9
               |||:
               APYI 82



               otein AAF98431.1 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C;Accession: E86408
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar,
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marzia
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719
A;Accession: E86408
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-581 <STO>
A;Cross-references: GB:AE005172; NID:g9795613; PIDN:AAF98431.1; GSPDB:GN00141
C;Genetics:
A;Map position: 1


  Query Match             74.5%;  Score 38;  DB 2;  Length 581;
  Best Local Similarity   87.5%;  Pred. No. 17;
  Matches    7;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 RMFPNAPY 8
           |||||| |
Db     349 RMFPNARY 356


RESULT   9
S42210
ISG-54K protein - Chinese hamster
C;Species: Cricetulus griseus (Chinese hamster)
C;Date: 06-Feb-1995 #sequence_revision 06-Feb-1995 #text_change 07-May-1999
C;Accession: S42210
R;Bluyssen, H.A.R.; Vlietstra, R.J.; van der Made, A.; Trapman, J.
Eur. J. Biochem. 220, 395-402, 1994
A;Title: The interferon-stimulated gene 54 K promoter contains two adjacent functiona
alpha inducibility.
A;Reference number: S42210; MUID:94170787
A;Accession: S42210
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-466 <BLU>
C;Superfamily: interferon-induced 56K protein


  Query Match             72.5%;  Score 37;  DB 2;  Length 466;
  Best Local Similarity   66.7%;  Pred. No. 21;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 RMFPNAPYL 9
           || ||:||:
Db     201 RMSPNSPYV 209


RESULT  10
S33558
unc-33 protein - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 07-Apr-1994 #sequence_revision 02-Jun-1994 #text_change 21-Jul-2000
C;Accession: S33558; S33559; S24643; S24644
R;Li, W.; Herman, R.K.; Shaw, J.E.
Genetics 132, 675-689, 1992
A;Title: Analysis of the Caenorhabditis elegans axonal guidance and outgrowth gene un
A;Reference number: S33558; MUID:93106371
A;Accession: S33558
A;Molecule type: DNA
A;Residues: 1-854 <LIW1>
A;Cross-references: EMBL:Z14148; NID:g6899; PIDN:CAA78520.1; PID:g6900
```

```
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 RMFPNAPYL 9
          |||||||||
Db    125 RMFPNAPYL 133


RESULT   3
A38080
Wilms tumor susceptibility protein WT1 - human
C;Species: Homo sapiens (man)
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 20-Jun-2000
C;Accession: A38080; S08273; A34673; I38504; I52811; I58315; A56411; S26286
R;Gessler, M.; Konig, A.; Bruns, G.A.
Genomics 12, 807-813, 1992
A;Title: The genomic organization and expression of the WT1 gene.
A;Reference number: A38080; MUID:92241883
A;Accession: A38080
A;Molecule type: DNA
A;Residues: 1-449 <GES1>
A;Cross-references: GB:X61631; GB:S99414; NID:g37981; PIDN:CAA43819.1; PID:g825731
A;Note: it is uncertain whether Met-1 is the initiator or whether translation is initiat
A;Note: sequence extracted from NCBI backbone (NCBIN:99414, NCBIN:99422, NCBIN:99479, NC
A;Note: the sequence in GenBank entry HSWTGEEX1, release 113.0, PIDN:CAA43819.1 differs
R;Gessler, M.; Poustka, A.; Cavenee, W.; Neve, R.L.; Orkin, S.H.; Bruns, G.A.P.
Nature 343, 774-778, 1990
A;Title: Homozygous deletion in Wilms tumours of a zinc-finger gene identified by chromo
A;Reference number: S08273; MUID:90158822
A;Accession: S08273
A;Molecule type: mRNA
A;Residues: 'SRQRPHPGALRNPTACPLPHFPPSLPPTHSPTHPPRAGTAAQAPGPRRLLAAILDFLLLQDPASTCVPEPASQHT
A;Cross-references: EMBL:X51630; NID:g37977; PIDN:CAA35956.1; PID:g37978
R;Call, K.M.; Glaser, T.; Ito, C.Y.; Buckler, A.J.; Pelletier, J.; Haber, D.A.; Rose, E.
Cell 60, 509-520, 1990
A;Title: Isolation and characterization of a zinc finger polypeptide gene at the human c
A;Reference number: A34673; MUID:90150277
A;Accession: A34673
A;Molecule type: mRNA
A;Residues: 85-249,267-364,'F',366-386,'T',388-407,411-449 <CAL>
A;Cross-references: GB:M30393; NID:g340381; PIDN:AAA36810.1; PID:g340382
R;Haber, D.A.; Sohn, R.L.; Buckler, A.J.; Pelletier, J.; Call, K.M.; Housman, D.E.
Proc. Natl. Acad. Sci. U.S.A. 88, 9618-9622, 1991
A;Title: Alternative splicing and genomic structure of the Wilms tumor gene WT1.
A;Reference number: A56411; MUID:92052142
A;Contents: annotation; alternative splicing
R;Phelan, S.A.; Lindberg, C.; Call, K.M.
Cell Growth Differ. 5, 677-686, 1994
A;Title: Wilms' tumor gene, WT1, mRNA is down-regulated during induction of erythroid ar
A;Reference number: I38504; MUID:94368704
A;Accession: I38504
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-18 <PHE>
A;Cross-references: EMBL:U06486; NID:g473563; PIDN:AAA62865.1; PID:g458432
R;Pelletier, J.; Bruening, W.; Kashtan, C.E.; Mauer, S.M.; Manivel, J.C.; Striegel, J.E.
Cell 67, 437-447, 1991
A;Title: Germline mutations in the Wilms' tumor suppressor gene are associated with abnc
A;Reference number: I52811; MUID:92005721
A;Accession: I52811
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 355-365,'H',367-377 <PEL>
A;Cross-references: GB:S61513; NID:g237599; PIDN:AAB20109.1; PID:g237600
A;Note: mutant form
R;Hamilton, T.B.; Barilla, K.C.; Romaniuk, P.J.
Nucleic Acids Res. 23, 277-284, 1995
A;Title: High affinity binding sites for the Wilms' tumour suppressor protein WT1.
A;Reference number: I58315; MUID:95166649
A;Accession: I58315
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 'MGHHHHHHHHHHSSGHIEGRHM',301-364,'F',366-386,'T',388-407,411-449 <HAM>
A;Cross-references: GB:S75264; NID:g896246; PIDN:AAB33443.1; PID:g896247
```

```
A;Note: this sequence is engineered
C;Genetics:
A;Gene: GDB:WT1
A;Cross-references: GDB:120496; OMIM:194070
A;Map position: 11p13-11p13
A;Introns: 148/1; 189/1; 223/2; 249/2; 266/2; 298/3; 349/1; 379/1; 410/1
A;Note: mRNA transcripts containing both alternatively spliced regions are the most a
C;Keywords: alternative splicing; DNA binding; kidney; tumor suppressor; zinc finger
F;1-449/Product: Wilms tumor susceptibility protein WT1, splice form 1 #status predic
F;1-407,411-449/Product: Wilms tumor susceptibility protein WT1, splice form 4 #statu
F;1-249,267-449/Product: Wilms tumor susceptibility protein WT1, splice form 3 #statu
F;1-249,267-407,411-449/Product: Wilms tumor susceptibility protein WT1, splice form


  Query Match             100.0%;  Score 51;  DB 2;  Length 449;
  Best Local Similarity   100.0%;  Pred. No. 0.042;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 RMFPNAPYL 9
          |||||||||
Db    126 RMFPNAPYL 134


RESULT   4
A39692
Wilms' tumor protein analog, WT1 - mouse
C;Species: Mus musculus (house mouse)
C;Date: 30-Dec-1991 #sequence_revision 30-Dec-1991 #text_change 16-Feb-1997
C;Accession: A39692
R;Buckler, A.J.; Pelletier, J.; Haber, D.A.; Glaser, T.; Housman, D.E.
Mol. Cell. Biol. 11, 1707-1712, 1991
A;Title: Isolation, characterization, and expression of the murine Wilms' tumor gene
A;Reference number: A39692; MUID:91141522
A;Accession: A39692
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-449 <BUC>
A;Cross-references: GB:M55512
C;Keywords: alternative splicing; DNA binding; transcription regulation; tumor suppre


  Query Match             100.0%;  Score 51;  DB 2;  Length 449;
  Best Local Similarity   100.0%;  Pred. No. 0.042;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 RMFPNAPYL 9
          |||||||||
Db    126 RMFPNAPYL 134


RESULT   5
G64144
hypothetical protein HI0177 - Haemophilus influenzae (strain Rd KW20)
C;Species: Haemophilus influenzae
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 28-Jul-2000
C;Accession: G64144
R;Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage
; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman
, D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M.
Science 269, 496-512, 1995
A;Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Vente
A;Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
A;Reference number: A64000; MUID:95350630
A;Accession: G64144
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-262 <TIGR>
A;Cross-references: GB:U32703; GB:L42023; NID:g1573133; PIDN:AAC21847.1; PID:g1573134
A;Note: best homolog was a hypothetical protein from Pseudomonas aeruginosa
C;Superfamily: conserved hypothetical protein HI0177


  Query Match              78.4%;  Score 40;  DB 1;  Length 262;
```

```
                   GenCore version 4.5
            Copyright (c) 1993 - 2000  Compugen Ltd.

OM protein - protein search, using sw model

Run on:          October 23, 2001, 13:20:07 ; Search time 27.18 Seconds
                                              (without alignments)
                                              25.223 Million cell updates/sec

Title:           US-09-625-963-1
Perfect score:   51
Sequence:        1 RMFPNAPYL 9

Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched:        219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters:      219241

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       PIR_68:*
                1:  pir1:*
                2:  pir2:*
                3:  pir3:*
                4:  pir4:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                               SUMMARIES
                     %
Result             Query
  No.    Score     Match Length DB  ID                           Description
-------------------------------------------------------------------------------
    1      51     100.0    410  2  JC5046                        Wilms' tumor suppr
    2      51     100.0    448  2  S33926                        Wilms' tumor prote
    3      51     100.0    449  2  A38080                        Wilms tumor suscep
    4      51     100.0    449  2  A39692                        Wilms' tumor prote
    5      40      78.4    262  1  G64144                        hypothetical prote
    6      39      76.5    203  2  C71375                        conserved hypothet
    7      38      74.5    205  2  F83495                        probable hydrolase
    8      38      74.5    581  2  E86408                        hypothetical prote
    9      37      72.5    466  2  S42210                        ISG-54K protein -
   10      37      72.5    854  2  S33558                        unc-33 protein - C
   11      37      72.5   1232  2  A55478                        neuronal apoptosis
   12      36      70.6     85  2  S17463                        glutathione transf
   13      36      70.6    115  2  S39393                        glutathione transf
   14      36      70.6    142  2  S17462                        glutathione transf
   15      36      70.6    217  2  JX0095                        glutathione transf
   16      36      70.6    218  1  XURTG4                        glutathione transf
   17      36      70.6    218  2  B34159                        glutathione transf
   18      36      70.6    218  2  A29036                        glutathione transf
   19      36      70.6    218  2  S13202                        glutathione transf
   20      36      70.6    218  2  S33860                        glutathione transf
   21      36      70.6    218  2  A46143                        mu-class glutathio
   22      36      70.6    218  2  A23732                        glutathione transf
   23      36      70.6    218  2  A39375                        glutathione transf
   24      36      70.6    218  2  S32425                        glutathione transf
   25      36      70.6    218  2  A46048                        glutathione transf
   26      36      70.6    218  2  S01719                        glutathione transf
   27      36      70.6    218  2  B28946                        glutathione transf
   28      36      70.6    218  2  A29794                        glutathione transf
   29      36      70.6    218  2  A47486                        glutathione transf
   30      36      70.6    218  2  S65674                        glutathione transf
   31      36      70.6    220  2  S18464                        glutathione transf
   32      36      70.6    225  2  A35295                        glutathione transf
   33      36      70.6    358  2  JC5964                        apoptosis inhibito
   34      36      70.6    381  2  S65212                        hypothetical prote
   35      36      70.6    382  2  A48492                        polysaccharide exp
   36      36      70.6    432  2  B96515                        hypothetical prote
   37      36      70.6    434  2  C96515                        hypothetical prote
   38      36      70.6    487  2  T07960                        probable (S)-N-met
   39      36      70.6    488  2  T07963                        probable (S)-N-met
   40      36      70.6    618  2  S68450                        apoptosis inhibito
   41      36      70.6   1249  2  T14150                        vesicle associated
   42      35      68.6    194  2  A82821                        hypothetical prote
   43      35      68.6    273  2  S40003                        trypsin-related pr
   44      35      68.6    345  2  A82348                        probable ADP-hepto
   45      35      68.6    380  2  T42755                        tyrosylprotein sul

                               ALIGNMENTS

RESULT   1
JC5046
Wilms' tumor suppressor protein - African clawed frog
N;Alternate names: WT1
C;Species: Xenopus laevis (African clawed frog)
C;Date: 31-Jan-1997 #sequence_revision 31-Jan-1997 #text_change 31-Jan-1997
C;Accession: JC5046
R;Semba, K.; Saito-Ueno, R.; Takayama, G.; Kondo, M.
Gene 175, 167-172, 1996
A;Title: cDNA cloning and its pronephros-specific expression of the Wilms' tumor supp
A;Reference number: JC5046; MUID:97074667
A;Contents: testis
A;Accession: JC5046
A;Molecule type: mRNA
A;Residues: 1-410 <SEM>
A;Cross-references: DDBJ:D82051
C;Comment: This protein is involved in kidney morphogenesis.
C;Genetics:
A;Gene: wt1

  Query Match             100.0%;  Score 51;  DB 2;  Length 410;
  Best Local Similarity   100.0%;  Pred. No. 0.038;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 RMFPNAPYL 9
            |||||||||
Db      107 RMFPNAPYL 115


RESULT   2
S33926
Wilms' tumor protein WT1 - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 10-May-1996
C;Accession: S33926
R;Sharma, P.M.; Yang, X.; Bowman, M.; Roberts, V.; Sukumar, S.
Cancer Res. 52, 6407-6412, 1992
A;Title: Molecular cloning of rat Wilms' tumor complementary DNA and a study of messe
A;Reference number: S33926; MUID:93046155
A;Accession: S33926
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-448 <SHA>
A;Cross-references: EMBL:X69716
C;Genetics:
A;Gene: WT1
C;Keywords: tumor suppressor

  Query Match             100.0%;  Score 51;  DB 2;  Length 448;
  Best Local Similarity   100.0%;  Pred. No. 0.042;
```

Frog

```
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/804,198
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  CANTRELL, PAUL R.
;      REGISTRATION NUMBER:  36,470
;      REFERENCE/DOCKET NUMBER:  P9113
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  317-276-3885
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1996 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  peptide
US-08-804-198-3


  Query Match             62.1%;  Score 36;  DB 2;  Length 1996;
  Best Local Similarity   71.4%;  Pred. No. 5.1e+02;
  Matches    5;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      3 MPFPGPL 9
          :|||||:
Db   1281 LPFPGPV 1287


RESULT  15
US-09-269-220-1
; Sequence 1, Application US/09269220
; Patent No. 6180761
; GENERAL INFORMATION:
;  APPLICANT:  HAN, Sang K
;  APPLICANT:  SHIN, Yoo C
;  TITLE OF INVENTION:  CASEIN AND PROCESS FOR THE PREPARATION THEREOF
;  FILE REFERENCE:  1423.1001/MJH
;  CURRENT APPLICATION NUMBER: US/09/269,220
;  CURRENT FILING DATE:  1999-03-23
;  PRIOR APPLICATION NUMBER:  KR 1996-43482
;  PRIOR FILING DATE:  1996-03-23
;  PRIOR APPLICATION NUMBER:  PCT/KR97/00182
;  PRIOR FILING DATE:  1997-09-23
;  NUMBER OF SEQ ID NOS:  7
;  SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 209
;   TYPE: PRT
;   ORGANISM: Bos taurus
;   FEATURE:
;   NAME/KEY: ACT_SITE
;   LOCATION: (15)
;   OTHER INFORMATION: phospholyated serine
;   NAME/KEY: ACT_SITE
;   LOCATION: (17)..(19)
US-09-269-220-1


  Query Match             60.3%;  Score 35;  DB 4;  Length 209;
  Best Local Similarity   83.3%;  Pred. No. 79;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      4 PFPGPL 9
          |||||:
Db     61 PFPGPI 66


Search completed: October 23, 2001, 13:21:33
Job time: 86 sec
```

```
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-728-626-5

  Query Match             63.8%;  Score 37;  DB 2;  Length 778;
  Best Local Similarity   66.7%;  Pred. No. 1.4e+02;
  Matches    6;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy        1 HLMPFPGPL 9
            |||| | |:
Db      303 HLMPSPAPV 311


RESULT  12
US-08-808-599A-5
; Sequence 5, Application US/08808599A
; Patent No. 6111089
;  GENERAL INFORMATION:
;    APPLICANT:  Fukuda, Michiko N.
;    TITLE OF INVENTION:  Trophinin, Trophinin-Assisting
;    TITLE OF INVENTION:  Proteins and Methods to Inhibit Implantation
;    NUMBER OF SEQUENCES:  41
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Campbell & Flores LLP
;      STREET:  4370 La Jolla Village Drive, Suite 700
;      CITY:  San Diego
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  92122
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/808,599A
;      FILING DATE:  28-FEB-1997
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/317,522
;      FILING DATE:  04-OCT-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/439,818
;      FILING DATE:  12-MAY-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Campbell, Cathryn A.
;      REGISTRATION NUMBER:  31,815
;      REFERENCE/DOCKET NUMBER:  P-LA 2256
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619) 535-9001
;      TELEFAX:  (619) 535-8949
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  778 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-808-599A-5

  Query Match             63.8%;  Score 37;  DB 3;  Length 778;
  Best Local Similarity   66.7%;  Pred. No. 1.4e+02;
  Matches    6;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy        1 HLMPFPGPL 9
            |||| | |:
Db      303 HLMPSPAPV 311


RESULT  13
US-08-804-227C-9
```

```
; Sequence 9, Application US/08804227C
; Patent No. 5876991
;  GENERAL INFORMATION:
;    APPLICANT:  DeHoff, Bradley S.
;    APPLICANT:  Kuhstoss, Stuart A.
;    APPLICANT:  Rosteck, Paul R., Jr.
;    APPLICANT:  Sutton, Kimberly L.
;    TITLE OF INVENTION:  POLYKETIDE SYNTHASE GENES
;    NUMBER OF SEQUENCES:  15
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  THOMAS G. PLANT  1501
;      STREET:  LILLY CORPORATE CENTER
;      CITY:  INDIANAPOLIS
;      STATE:  IN
;      COUNTRY:  USA
;      ZIP:  46285
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  ASCI(DOS) Text only
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/804,227C
;      FILING DATE:  February 21, 1997
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Plant, Thomas, G.
;      REGISTRATION NUMBER:  35,784
;      REFERENCE/DOCKET NUMBER:  X-8231
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  317-276-2459
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1996 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  peptide
US-08-804-227C-9

  Query Match             62.1%;  Score 36;  DB 2;  Length 1996;
  Best Local Similarity   71.4%;  Pred. No. 5.1e+02;
  Matches    5;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        3 MPFPGPL 9
            :|||||:
Db     1281 LPFPGPV 1287


RESULT  14
US-08-804-198-3
; Sequence 3, Application US/08804198
; Patent No. 5945320
;  GENERAL INFORMATION:
;    APPLICANT:  Burgett, Stanley G.
;    APPLICANT:  Kuhstoss, Stuart A.
;    APPLICANT:  Rao, Nagaraja R.
;    APPLICANT:  Richardson, Mark A.
;    APPLICANT:  Rosteck, Paul R., Jr.
;    TITLE OF INVENTION:  PLATENOLIDE SYNTHASE GENE
;    NUMBER OF SEQUENCES:  6
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  PAUL R. CANTRELL  1138
;      STREET:  LILLY CORPORATE CENTER
;      CITY:  INDIANAPOLIS
;      STATE:  IN
;      COUNTRY:  USA
;      ZIP:  46285
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  Macintosh
;      OPERATING SYSTEM:  Macintosh 7.0
;      SOFTWARE:  Microsoft Word 5.1
```

```
RESULT   9
US-08-751-965-5
; Sequence 5, Application US/08751965
; Patent No. 5858360
;  GENERAL INFORMATION:
;    APPLICANT:  Fukuda, Michiko N.
;    TITLE OF INVENTION:  Trophinin and Trophinin-Assisting
;    TITLE OF INVENTION:  Proteins
;    NUMBER OF SEQUENCES:  22
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Campbell and Flores
;      STREET:  4370 La Jolla Village Drive, Suite 700
;      CITY:  San Diego
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  92122
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/751,965
;      FILING DATE:  Herewith
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/439,818
;      FILING DATE:  12-MAY-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Campbell, Cathryn A.
;      REGISTRATION NUMBER:  31,815
;      REFERENCE/DOCKET NUMBER:  P-LA 2252
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619) 535-9001
;      TELEFAX:  (619) 535-8949
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  778 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-751-965-5


  Query Match             63.8%;  Score 37;  DB 2;  Length 778;
  Best Local Similarity   66.7%;  Pred. No. 1.4e+02;
  Matches    6;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 HLMPFPGPL 9
           |||| | |:
Db     303 HLMPSPAPV 311


RESULT  10
US-08-738-975-5
; Sequence 5, Application US/08738975
; Patent No. 5880267
;  GENERAL INFORMATION:
;    APPLICANT:  Fukuda, Michiko N.
;    TITLE OF INVENTION:  Trophinin and Trophinin-Assisting
;    TITLE OF INVENTION:  Proteins
;    NUMBER OF SEQUENCES:  22
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Campbell and Flores
;      STREET:  4370 La Jolla Village Drive, Suite 700
;      CITY:  San Diego
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  92122
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/738,975
;      FILING DATE:  herewith
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/439,818
;      FILING DATE:  05-Dec-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Campbell, Cathryn A.
;      REGISTRATION NUMBER:  31,815
;      REFERENCE/DOCKET NUMBER:  P-LA 2251
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619) 535-9001
;      TELEFAX:  (619) 535-8949
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  778 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-738-975-5


  Query Match             63.8%;  Score 37;  DB 2;  Length 778;
  Best Local Similarity   66.7%;  Pred. No. 1.4e+02;
  Matches    6;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 HLMPFPGPL 9
           |||| | |:
Db     303 HLMPSPAPV 311


RESULT  11
US-08-728-626-5
; Sequence 5, Application US/08728626
; Patent No. 5910451
;  GENERAL INFORMATION:
;    APPLICANT:  Fukuda, Michiko N.
;    TITLE OF INVENTION:  Trophinin and Trophinin-Assisting
;    TITLE OF INVENTION:  Proteins
;    NUMBER OF SEQUENCES:  22
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Campbell and Flores
;      STREET:  4370 La Jolla Village Drive, Suite 700
;      CITY:  San Diego
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  92122
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/728,626
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/439,818
;      FILING DATE:  12-MAY-1995
;      APPLICATION NUMBER:  US 08/317,522
;      FILING DATE:  04-OCT-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Campbell, Cathryn A.
;      REGISTRATION NUMBER:  31,815
;      REFERENCE/DOCKET NUMBER:  P-LA 1563
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619) 535-9001
;      TELEFAX:  (619) 535-8949
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  778 amino acids
```

```
;   ORGANISM: Sarcophaga peregrina
US-09-120-365-1

  Query Match             65.5%;  Score 38;  DB 3;  Length 550;
  Best Local Similarity   100.0%;  Pred. No. 67;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      3 MPFPGP 8
          ||||||
Db    214 MPFPGP 219


RESULT   6
US-09-515-039-1
; Sequence 1, Application US/09515039
; Patent No. 6214599
; GENERAL INFORMATION:
;  APPLICANT: Natori, Shunji
;  TITLE OF INVENTION: NEW PROTEASE
;  FILE REFERENCE: 32290-144749
;  CURRENT APPLICATION NUMBER: US/09/515,039
;  CURRENT FILING DATE: 2000-03-06
;  EARLIER APPLICATION NUMBER: JP 9-333 474
;  EARLIER FILING DATE: 1997-11-18
;  NUMBER OF SEQ ID NOS: 101
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
;   LENGTH: 550
;   TYPE: PRT
;   ORGANISM: Sarcophaga peregrina
US-09-515-039-1


  Query Match             65.5%;  Score 38;  DB 4;  Length 550;
  Best Local Similarity   100.0%;  Pred. No. 67;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      3 MPFPGP 8
          ||||||
Db    214 MPFPGP 219


RESULT   7
US-08-317-522A-5
; Sequence 5, Application US/08317522A
; Patent No. 5599918
;  GENERAL INFORMATION:
;    APPLICANT:  Fukuda, Michiko N.
;    TITLE OF INVENTION:  Trophinin and Trophinin-Assisting
;    TITLE OF INVENTION:  Proteins
;    NUMBER OF SEQUENCES:  13
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Campbell and Flores
;      STREET:  4370 La Jolla Village Drive, Suite 700
;      CITY:  San Diego
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  92122
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/317,522A
;      FILING DATE:  04-OCT-1994
;      CLASSIFICATION:  536
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Campbell, Cathryn A.
;      REGISTRATION NUMBER:  31,815
;      REFERENCE/DOCKET NUMBER:  P-LA 9991
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619) 535-9001
;      TELEFAX:  (619) 535-8949
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  732 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-317-522A-5

  Query Match             63.8%;  Score 37;  DB 1;  Length 732;
  Best Local Similarity   66.7%;  Pred. No. 1.3e+02;
  Matches    6;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 HLMPFPGPL 9
          |||| | |:
Db    303 HLMPSPAPV 311


RESULT   8
US-08-439-818A-5
; Sequence 5, Application US/08439818A
; Patent No. 5654145
;  GENERAL INFORMATION:
;    APPLICANT:  Fukuda, Michiko N.
;    TITLE OF INVENTION:  Trophinin and Trophinin-Assisting
;    TITLE OF INVENTION:  Proteins
;    NUMBER OF SEQUENCES:  22
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Campbell and Flores
;      STREET:  4370 La Jolla Village Drive, Suite 700
;      CITY:  San Diego
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  92122
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/439,818A
;      FILING DATE:  12-MAY-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/317,522
;      FILING DATE:  04-OCT-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Campbell, Cathryn A.
;      REGISTRATION NUMBER:  31,815
;      REFERENCE/DOCKET NUMBER:  P-LA 1563
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619) 535-9001
;      TELEFAX:  (619) 535-8949
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  778 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-439-818A-5


  Query Match             63.8%;  Score 37;  DB 1;  Length 778;
  Best Local Similarity   66.7%;  Pred. No. 1.4e+02;
  Matches    6;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 HLMPFPGPL 9
          |||| | |:
Db    303 HLMPSPAPV 311
```

```
  Query Match             65.5%;  Score 38;  DB 4;  Length 50;
  Best Local Similarity   70.0%;  Pred. No. 6.3;
  Matches    7;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 HLMPFPGPLL 10
              ||  |||| |
Db         20 HLASFPGPSL 29


RESULT   2
US-09-120-365-60
; Sequence 60, Application US/09120365
; Patent No. 6103514
; GENERAL INFORMATION:
;  APPLICANT: Natori, Shunji
;  TITLE OF INVENTION: NEW PROTEASE
;  FILE REFERENCE: 32290-144749
;  CURRENT APPLICATION NUMBER: US/09/120,365
;  CURRENT FILING DATE: 1998-07-22
;  EARLIER APPLICATION NUMBER: JP 9-333 474
;  EARLIER FILING DATE: 1997-11-18
;  NUMBER OF SEQ ID NOS: 101
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 60
;   LENGTH: 340
;   TYPE: PRT
;   ORGANISM: Sarcophaga peregrina
US-09-120-365-60


  Query Match             65.5%;  Score 38;  DB 3;  Length 340;
  Best Local Similarity   100.0%;  Pred. No. 42;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 MPFPGP 8
              ||||||
Db          4 MPFPGP 9


RESULT   3
US-09-515-039-60
; Sequence 60, Application US/09515039
; Patent No. 6214599
; GENERAL INFORMATION:
;  APPLICANT: Natori, Shunji
;  TITLE OF INVENTION: NEW PROTEASE
;  FILE REFERENCE: 32290-144749
;  CURRENT APPLICATION NUMBER: US/09/515,039
;  CURRENT FILING DATE: 2000-03-06
;  EARLIER APPLICATION NUMBER: JP 9-333 474
;  EARLIER FILING DATE: 1997-11-18
;  NUMBER OF SEQ ID NOS: 101
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 60
;   LENGTH: 340
;   TYPE: PRT
;   ORGANISM: Sarcophaga peregrina
US-09-515-039-60


  Query Match             65.5%;  Score 38;  DB 4;  Length 340;
  Best Local Similarity   100.0%;  Pred. No. 42;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 MPFPGP 8
              ||||||
Db          4 MPFPGP 9


RESULT   4
US-08-702-598-2
; Sequence 2, Application US/08702598
; Patent No. 5880332
;  GENERAL INFORMATION:
;    APPLICANT:  Camara, Bilal
;    APPLICANT:  Kuntz, Marcel
;    TITLE OF INVENTION:  DNA Constructs, Cells and Plants Derived
;    TITLE OF INVENTION:  Therefrom
;    NUMBER OF SEQUENCES:  2
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Spencer & Frank
;      STREET:  1100 New York Avenue, N.W. Suite 300E
;      CITY:  Washington, D.C.
;      COUNTRY:  USA
;      ZIP:  20005
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/702,598
;      FILING DATE:
;      CLASSIFICATION:  800
;      CLASSIFICATION:  C12N 15/82
;      CLASSIFICATION:  C12N 9/00
;      CLASSIFICATION:  A01H 5/00
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  GB 9403943.5
;      FILING DATE:  01-MAR-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  EP 94400626.1
;      FILING DATE:  23-MAR-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Calvetti, Frecerick F.
;      REGISTRATION NUMBER:  28,557
;      REFERENCE/DOCKET NUMBER:  GROFO 7000
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 414-4000
;      TELEFAX:  (202) 414-4040
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  498 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-702-598-2


  Query Match             65.5%;  Score 38;  DB 2;  Length 498;
  Best Local Similarity   77.8%;  Pred. No. 61;
  Matches    7;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 LMPFPGPLL 10
              | ||| |||
Db          5 LKPFPSPLL 13


RESULT   5
US-09-120-365-1
; Sequence 1, Application US/09120365
; Patent No. 6103514
; GENERAL INFORMATION:
;  APPLICANT: Natori, Shunji
;  TITLE OF INVENTION: NEW PROTEASE
;  FILE REFERENCE: 32290-144749
;  CURRENT APPLICATION NUMBER: US/09/120,365
;  CURRENT FILING DATE: 1998-07-22
;  EARLIER APPLICATION NUMBER: JP 9-333 474
;  EARLIER FILING DATE: 1997-11-18
;  NUMBER OF SEQ ID NOS: 101
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
;   LENGTH: 550
;   TYPE: PRT
```

```
                    GenCore version 4.5
         Copyright (c) 1993 - 2000  Compugen Ltd.


OM protein - protein search, using sw model

Run on:         October 23, 2001, 13:21:32 ; Search time 23.71 Seconds
                                             (without alignments)
                                             8.684 Million cell updates/sec

Title:          US-09-625-963-3
Perfect score:  58
Sequence:       1 HLMPFPGPLL 10

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched:       197339 seqs, 20590346 residues

Total number of hits satisfying chosen parameters:     197339

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       Issued_Patents_AA:*
               1:  /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
               2:  /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
               3:  /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
               4:  /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
               5:  /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
               6:  /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                               SUMMARIES
               %
Result       Query
  No.  Score  Match Length DB  ID                          Description
-----------------------------------------------------------------------
    1     38   65.5     50  4  US-08-905-223-378            Sequence 378, App
    2     38   65.5    340  3  US-09-120-365-60             Sequence 60, Appl
    3     38   65.5    340  4  US-09-515-039-60             Sequence 60, Appl
    4     38   65.5    498  2  US-08-702-598-2              Sequence 2, Appli
    5     38   65.5    550  3  US-09-120-365-1              Sequence 1, Appli
    6     38   65.5    550  4  US-09-515-039-1              Sequence 1, Appli
    7     37   63.8    732  1  US-08-317-522A-5             Sequence 5, Appli
    8     37   63.8    778  1  US-08-439-818A-5             Sequence 5, Appli
    9     37   63.8    778  2  US-08-751-965-5              Sequence 5, Appli
   10     37   63.8    778  2  US-08-738-975-5              Sequence 5, Appli
   11     37   63.8    778  2  US-08-728-626-5              Sequence 5, Appli
   12     37   63.8    778  3  US-08-808-599A-5             Sequence 5, Appli
   13     36   62.1   1996  2  US-08-804-227C-9             Sequence 9, Appli
   14     36   62.1   1996  2  US-08-804-198-3              Sequence 3, Appli
   15     35   60.3    209  4  US-09-269-220-1              Sequence 1, Appli
   16     35   60.3    209  4  US-09-269-220-2              Sequence 2, Appli
   17     35   60.3    230  3  US-09-320-878-17             Sequence 17, Appl
   18     35   60.3    518  2  US-08-836-620A-18            Sequence 18, Appl
   19     35   60.3    721  4  US-09-146-221-4              Sequence 4, Appli
   20     35   60.3    724  4  US-09-146-221-6              Sequence 6, Appli
   21     35   60.3    727  4  US-09-146-221-8              Sequence 8, Appli
   22     34   58.6     13  2  US-08-747-137-86             Sequence 86, Appl
   23     34   58.6    377  2  US-08-969-106-6              Sequence 6, Appli
   24     34   58.6    377  4  US-09-054-492B-1             Sequence 1, Appli
   25     34   58.6    503  4  US-09-147-009-11             Sequence 11, Appl
   26     34   58.6   2214  1  US-08-727-034-7              Sequence 7, Appli
   27     33   56.9      9  1  US-08-463-620-8              Sequence 8, Appli
   28     33   56.9      9  2  US-08-224-917-8              Sequence 8, Appli
   29     33   56.9      9  2  US-08-914-853-8              Sequence 8, Appli
   30     33   56.9      9  5  PCT-US95-03934A-8            Sequence 8, Appl
   31     33   56.9     78  1  US-08-487-359-5              Sequence 5, Appli
   32     33   56.9     78  1  US-08-222-798A-5             Sequence 5, Appli
   33     33   56.9    192  3  US-08-486-099-107            Sequence 107, App
   34     33   56.9    192  3  US-08-360-107A-117           Sequence 117, App
   35     33   56.9    192  3  US-08-484-223B-107           Sequence 107, App
   36     33   56.9    192  3  US-08-919-597-107            Sequence 107, App
   37     33   56.9    192  3  US-08-475-668A-107           Sequence 107, App
   38     33   56.9    192  3  US-08-485-551A-107           Sequence 107, App
   39     33   56.9    192  3  US-08-471-913A-107           Sequence 107, App
   40     33   56.9    192  4  US-08-485-264A-107           Sequence 107, App
   41     33   56.9    277  2  US-08-403-852D-18            Sequence 18, Appl
   42     33   56.9    277  3  US-08-510-646B-19            Sequence 19, Appl
   43     33   56.9    277  4  US-09-231-818-18             Sequence 18, Appl
   44     33   56.9    335  1  US-08-347-826A-1             Sequence 1, Appli
   45     33   56.9    389  2  US-08-485-449-2              Sequence 2, Appli


                               ALIGNMENTS


RESULT  1
US-08-905-223-378
; Sequence 378, Application US/08905223
; Patent No. 6222029
;  GENERAL INFORMATION:
;    APPLICANT:  Edwards, Jean-Baptiste D.
;    APPLICANT:  Duelert, Aymeric
;    APPLICANT:  Lacroix, Bruno
;    TITLE OF INVENTION:  5' ESTs FOR SECRETED PROTEINS
;    NUMBER OF SEQUENCES:  503
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Knobbe, Martens, Olson & Bear
;      STREET:  501 West Broadway
;      CITY:  San Diego
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  92101-3505
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy Disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Win95
;      SOFTWARE:  Word
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/905,223
;      FILING DATE:
;      CLASSIFICATION:  536
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Israelsen, Ned A.
;      REGISTRATION NUMBER:  29,655
;      REFERENCE/DOCKET NUMBER:
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619) 235-8550
;      TELEFAX:  (619) 235-0176
;  INFORMATION FOR SEQ ID NO:  378:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  50 amino acids
;      TYPE:  AMINO ACID
;      TOPOLOGY:  LINEAR
;    MOLECULE TYPE:  PROTEIN
;    ORIGINAL SOURCE:
;      ORGANISM:  Homo Sapiens
;      TISSUE TYPE:  Brain
;    FEATURE:
;      NAME/KEY:  sig_peptide
;      LOCATION:  -37..-1
;      IDENTIFICATION METHOD:  Von Heijne matrix
;      OTHER INFORMATION:  score 3.9
;      OTHER INFORMATION:  seq PGPSLRLFSGSQA/SV
US-08-905-223-378
```

```
;   LOCATION: (82)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-758-462-903

  Query Match             63.8%;  Score 37;  DB 5;  Length 88;
  Best Local Similarity   100.0%;  Pred. No. 99;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      4 PFPGPL 9
          ||||||
Db     22 PFPGPL 27


Search completed: October 23, 2001, 13:22:51
Job time: 164 sec
```

```
;  PRIOR APPLICATION NUMBER: US 60/124,142
;  PRIOR FILING DATE: 1999-03-12
;  PRIOR APPLICATION NUMBER: US 60/138,597
;  PRIOR FILING DATE: 1999-06-11
;  PRIOR APPLICATION NUMBER: US 60/168,666
;  PRIOR FILING DATE: 1999-12-03
;  PRIOR APPLICATION NUMBER: US 60/125,359
;  PRIOR FILING DATE: 1999-03-19
;  PRIOR APPLICATION NUMBER: US 60/168,664
;  PRIOR FILING DATE: 1999-12-03
;  PRIOR APPLICATION NUMBER: US 60/126,051
;  PRIOR FILING DATE: 1999-03-23
;  PRIOR APPLICATION NUMBER: US 60/169,906

  Query Match             65.5%;  Score 38;  DB 5;  Length 117;
  Best Local Similarity   70.0%;  Pred. No. 91;
  Matches    7;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 HLMPFPGPLL 10
          ||:  |||||
Db     63 HLLTRPGPLL 72

RESULT 12
US-09-617-682A-3093
; Sequence 3093, Application US/09617682A
; GENERAL INFORMATION:
;  APPLICANT: ALEXANDROV, Nickolai
;  APPLICANT:  BROVER, Vyacheslav
;  TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
;  TITLE OF INVENTION:  Thereby
;  FILE REFERENCE: 2750-1063P
;  CURRENT APPLICATION NUMBER: US/09/617,682A
;  CURRENT FILING DATE:  2000-07-19
;  NUMBER OF SEQ ID NOS: 16871
; SEQ ID NO 3093
;   LENGTH: 161
;   TYPE: PRT
;   ORGANISM: Arabidopsis thaliana
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 1..161
;   OTHER INFORMATION: Xaa is any amino acid
;   NAME/KEY: misc_feature
;   LOCATION: 1..161
;   OTHER INFORMATION: Ceres Seq. ID 1348465
US-09-617-682A-3093

  Query Match             65.5%;  Score 38;  DB 5;  Length 161;
  Best Local Similarity   77.8%;  Pred. No. 1.2e+02;
  Matches    7;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      2 LMPFPGPLL 10
          |:||| |||
Db     34 LVPFPPPLL 42

RESULT 13
US-60-312-544-6518
; Sequence 6518, Application US/60312544
; GENERAL INFORMATION:
;  APPLICANT: Cao, Yongwei
;  APPLICANT:  Edgerton, Michael D
;  APPLICANT:  Hinkle, Gregory J.
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Liu, Jingdong
;  APPLICANT:  Stein, Joshua
;  TITLE OF INVENTION: cDNA SEQUENCES AND USES FOR PLANT IMPROVEMENT
;  FILE REFERENCE: 38-10(52726)A
;  CURRENT APPLICATION NUMBER: US/60/312,544
;  CURRENT FILING DATE:  2001-08-15
;  NUMBER OF SEQ ID NOS: 10730
; SEQ ID NO 6518
;   LENGTH: 450
;   TYPE: PRT
;   ORGANISM: Glycine max
;   FEATURE:
;   OTHER INFORMATION: Clone ID: 700748156_FLI
US-60-312-544-6518

  Query Match             65.5%;  Score 38;  DB 6;  Length 450;
  Best Local Similarity   85.7%;  Pred. No. 3.1e+02;
  Matches    6;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      3 MPFPGPL 9
          :||||||
Db    276 IPFPGPL 282

RESULT 14
PCT-US01-08656-9992
; Sequence 9992, Application PC/TUS0108656
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc
;  TITLE OF INVENTION: NOVEL MACROPHAGE NUCLEIC ACIDS AND POLYPEPTIDES
;  FILE REFERENCE: 21272-066
;  CURRENT APPLICATION NUMBER: PCT/US01/08656
;  CURRENT FILING DATE:  2001-04-16
;  PRIOR APPLICATION NUMBER: 09/522,929
;  PRIOR FILING DATE: 2000-04-18
;  PRIOR APPLICATION NUMBER: 09/770,160
;  PRIOR FILING DATE: 2001-01-26
;  NUMBER OF SEQ ID NOS: 10994
;  SOFTWARE: Custom
; SEQ ID NO 9992
;   LENGTH: 47
;   TYPE: PRT
;   ORGANISM: Homo sapiens
PCT-US01-08656-9992

  Query Match             63.8%;  Score 37;  DB 1;  Length 47;
  Best Local Similarity   62.5%;  Pred. No. 56;
  Matches    5;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 HLMPFPGP 8
          | :| |||
Db     25 HFLPLPGP 32

RESULT 15
US-09-758-462-903
; Sequence 903, Application US/09758462
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PM007
;  CURRENT APPLICATION NUMBER: US/09/758,462
;  CURRENT FILING DATE:  2001-01-11
;  PRIOR APPLICATION NUMBER: 60/179,065
;  PRIOR FILING DATE: 2000-01-31
;  PRIOR APPLICATION NUMBER: 60/180,628
;  PRIOR FILING DATE: 2000-02-04
;  NUMBER OF SEQ ID NOS: 1734
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 903
;   LENGTH: 88
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (74)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;   NAME/KEY: SITE
```

```
;  PRIOR APPLICATION NUMBER: PCT/US00/06782
;  PRIOR FILING DATE: 2000-03-16
;  PRIOR APPLICATION NUMBER: PCT/US00/06822
;  PRIOR FILING DATE: 2000-03-16
;  PRIOR APPLICATION NUMBER: PCT/US00/06791
;  PRIOR FILING DATE: 2000-03-16
;  PRIOR APPLICATION NUMBER: PCT/US00/06828
;  PRIOR FILING DATE: 2000-03-16
;  PRIOR APPLICATION NUMBER: PCT/US00/06823
;  PRIOR FILING DATE: 2000-03-16
;  PRIOR APPLICATION NUMBER: PCT/US00/06781
;  PRIOR FILING DATE: 2000-03-16
;  PRIOR APPLICATION NUMBER: PCT/US00/07505
;  PRIOR FILING DATE: 2000-03-22
;  PRIOR APPLICATION NUMBER: PCT/US00/07440
;  PRIOR FILING DATE: 2000-03-22
;  PRIOR APPLICATION NUMBER: PCT/US00/07506
;  PRIOR FILING DATE: 2000-03-22
;  PRIOR APPLICATION NUMBER: PCT/US00/07507
;  PRIOR FILING DATE: 2000-03-22
;  PRIOR APPLICATION NUMBER: PCT/US00/07535
;  PRIOR FILING DATE: 2000-03-22
;  PRIOR APPLICATION NUMBER: PCT/US00/07525
;  PRIOR FILING DATE: 2000-03-22
;  PRIOR APPLICATION NUMBER: PCT/US00/07534
;  PRIOR FILING DATE: 2000-03-22
;  PRIOR APPLICATION NUMBER: PCT/US00/07483
;  PRIOR FILING DATE: 2000-03-22
;  PRIOR APPLICATION NUMBER: PCT/US00/07526
;  PRIOR FILING DATE: 2000-03-22
;  PRIOR APPLICATION NUMBER: PCT/US00/07527
;  PRIOR FILING DATE: 2000-03-22
;  PRIOR APPLICATION NUMBER: PCT/US00/07661
;  PRIOR FILING DATE: 2000-03-23
;  PRIOR APPLICATION NUMBER: PCT/US00/07579
;  PRIOR FILING DATE: 2000-03-23
;  PRIOR APPLICATION NUMBER: PCT/US00/07723
;  PRIOR FILING DATE: 2000-03-23
;  PRIOR APPLICATION NUMBER: PCT/US00/07724
;  PRIOR FILING DATE: 2000-03-23
;  PRIOR APPLICATION NUMBER: PCT/US00/14929
;  PRIOR FILING DATE: 2000-06-01
;  PRIOR APPLICATION NUMBER: PCT/US00/07722
;  PRIOR FILING DATE: 2000-03-23
;  PRIOR APPLICATION NUMBER: PCT/US00/07578
;  PRIOR FILING DATE: 2000-03-23
;  PRIOR APPLICATION NUMBER: PCT/US00/07726
;  PRIOR FILING DATE: 2000-03-23
;  PRIOR APPLICATION NUMBER: PCT/US00/07677
;  PRIOR FILING DATE: 2000-03-23
;  PRIOR APPLICATION NUMBER: PCT/US00/07725
;  PRIOR FILING DATE: 2000-03-23
;  PRIOR APPLICATION NUMBER: PCT/US00/09070
;  PRIOR FILING DATE: 2000-04-06
;  PRIOR APPLICATION NUMBER: PCT/US00/08982
;  PRIOR FILING DATE: 2000-04-06
;  PRIOR APPLICATION NUMBER: PCT/US00/08983
;  PRIOR FILING DATE: 2000-04-06
;  PRIOR APPLICATION NUMBER: PCT/US00/09067
;  PRIOR FILING DATE: 2000-04-06
;  PRIOR APPLICATION NUMBER: PCT/US00/09066
;  PRIOR FILING DATE: 2000-04-06
;  PRIOR APPLICATION NUMBER: PCT/US00/09068
;  PRIOR FILING DATE: 2000-04-06
;  PRIOR APPLICATION NUMBER: PCT/US00/08981
;  PRIOR FILING DATE: 2000-04-06
;  PRIOR APPLICATION NUMBER: PCT/US00/08980
;  PRIOR FILING DATE: 2000-04-06
;  PRIOR APPLICATION NUMBER: PCT/US00/09071
;  PRIOR FILING DATE: 2000-04-06
;  PRIOR APPLICATION NUMBER: PCT/US00/09069
;  PRIOR FILING DATE: 2000-04-06
;  PRIOR APPLICATION NUMBER: PCT/US00/15136
;  PRIOR FILING DATE: 2000-06-01
;  PRIOR APPLICATION NUMBER: PCT/US00/14926
;  PRIOR FILING DATE: 2000-06-01
;  PRIOR APPLICATION NUMBER: PCT/US00/14963
;  PRIOR FILING DATE: 2000-06-01
;  PRIOR APPLICATION NUMBER: PCT/US00/15135
;  PRIOR FILING DATE: 2000-06-01
;  PRIOR APPLICATION NUMBER: PCT/US00/14934
;  PRIOR FILING DATE: 2000-06-01
;  PRIOR APPLICATION NUMBER: PCT/US00/14933
;  PRIOR FILING DATE: 2000-06-01
;  PRIOR APPLICATION NUMBER: PCT/US00/15137
;  PRIOR FILING DATE: 2000-06-01
;  PRIOR APPLICATION NUMBER: PCT/US00/14928
;  PRIOR FILING DATE: 2000-06-01
;  PRIOR APPLICATION NUMBER: PCT/US00/14973
;  PRIOR FILING DATE: 2000-06-01
;  PRIOR APPLICATION NUMBER: PCT/US00/14964
;  PRIOR FILING DATE: 2000-06-01
;  PRIOR APPLICATION NUMBER: PCT/US00/26376
;  PRIOR FILING DATE: 2000-09-26
;  PRIOR APPLICATION NUMBER: PCT/US00/26371
;  PRIOR FILING DATE: 2000-09-26
;  PRIOR APPLICATION NUMBER: PCT/US00/26324
;  PRIOR FILING DATE: 2000-09-26
;  PRIOR APPLICATION NUMBER: PCT/US00/26323
;  PRIOR FILING DATE: 2000-09-26
;  PRIOR APPLICATION NUMBER: PCT/US00/26337
;  PRIOR FILING DATE: 2000-09-26
;  PRIOR APPLICATION NUMBER: PCT/US01/13318
;  PRIOR FILING DATE: 2001-04-27
;  PRIOR APPLICATION NUMBER: US 60/124,146
;  PRIOR FILING DATE: 1999-03-12
;  PRIOR APPLICATION NUMBER: US 60/167,061
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/124,093
;  PRIOR FILING DATE: 1999-03-12
;  PRIOR APPLICATION NUMBER: US 60/166,989
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/124,145
;  PRIOR FILING DATE: 1999-03-12
;  PRIOR APPLICATION NUMBER: US 60/168,654
;  PRIOR FILING DATE: 1999-12-03
;  PRIOR APPLICATION NUMBER: US 60/124,099
;  PRIOR FILING DATE: 1999-03-12
;  PRIOR APPLICATION NUMBER: US 60/168,661
;  PRIOR FILING DATE: 1999-12-03
;  PRIOR APPLICATION NUMBER: US 60/124,096
;  PRIOR FILING DATE: 1999-03-12
;  PRIOR APPLICATION NUMBER: US 60/168,622
;  PRIOR FILING DATE: 1999-12-03
;  PRIOR APPLICATION NUMBER: US 60/124,143
;  PRIOR FILING DATE: 1999-03-12
;  PRIOR APPLICATION NUMBER: US 60/168,663
;  PRIOR FILING DATE: 1999-12-03
;  PRIOR APPLICATION NUMBER: US 60/124,095
;  PRIOR FILING DATE: 1999-03-12
;  PRIOR APPLICATION NUMBER: US 60/138,598
;  PRIOR FILING DATE: 1999-06-11
;  PRIOR APPLICATION NUMBER: US 60/168,665
;  PRIOR FILING DATE: 1999-12-03
;  PRIOR APPLICATION NUMBER: US 60/125,360
;  PRIOR FILING DATE: 1999-03-19
;  PRIOR APPLICATION NUMBER: US 60/138,626
;  PRIOR FILING DATE: 1999-06-11
;  PRIOR APPLICATION NUMBER: US 60/168,662
;  PRIOR FILING DATE: 1999-12-03
;  PRIOR APPLICATION NUMBER: US 60/124,144
;  PRIOR FILING DATE: 1999-03-12
;  PRIOR APPLICATION NUMBER: US 60/138,574
;  PRIOR FILING DATE: 1999-06-11
;  PRIOR APPLICATION NUMBER: US 60/168,667
;  PRIOR FILING DATE: 1999-12-03
```

```
;  PRIOR APPLICATION NUMBER: 60/239,937
;  PRIOR FILING DATE: 2000-10-13
;  PRIOR APPLICATION NUMBER: 60/241,787
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/246,474
;  PRIOR FILING DATE: 2000-11-08
;  PRIOR APPLICATION NUMBER: 60/246,532
;  PRIOR FILING DATE: 2000-11-08
;  PRIOR APPLICATION NUMBER: 60/249,216
;  PRIOR FILING DATE: 2000-11-17
;  PRIOR APPLICATION NUMBER: 60/249,210
;  PRIOR FILING DATE: 2000-11-17
;  PRIOR APPLICATION NUMBER: 60/226,681
;  PRIOR FILING DATE: 2000-08-22
;  PRIOR APPLICATION NUMBER: 60/225,759
;  PRIOR FILING DATE: 2000-08-14
;  PRIOR APPLICATION NUMBER: 60/225,213
;  PRIOR FILING DATE: 2000-08-14
;  PRIOR APPLICATION NUMBER: 60/227,182
;  PRIOR FILING DATE: 2000-08-22
;  PRIOR APPLICATION NUMBER: 60/225,214
;  PRIOR FILING DATE: 2000-08-14
;  PRIOR APPLICATION NUMBER: 60/235,836
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: 60/230,438
;  PRIOR FILING DATE: 2000-09-06
;  PRIOR APPLICATION NUMBER: 60/215,135
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: 60/225,266
;  PRIOR FILING DATE: 2000-08-14
;  PRIOR APPLICATION NUMBER: 60/249,218
;  PRIOR FILING DATE: 2000-11-17
;  PRIOR APPLICATION NUMBER: 60/249,208
;  PRIOR FILING DATE: 2000-11-17
;  PRIOR APPLICATION NUMBER: 60/249,213
;  PRIOR FILING DATE: 2000-11-17
;  PRIOR APPLICATION NUMBER: 60/249,212
;  PRIOR FILING DATE: 2000-11-17
;  PRIOR APPLICATION NUMBER: 60/249,207
;  PRIOR FILING DATE: 2000-11-17
;  PRIOR APPLICATION NUMBER: 60/249,245
;  PRIOR FILING DATE: 2000-11-17
;  PRIOR APPLICATION NUMBER: 60/249,244
;  PRIOR FILING DATE: 2000-11-17
;  PRIOR APPLICATION NUMBER: 60/249,217
;  PRIOR FILING DATE: 2000-11-17
;  PRIOR APPLICATION NUMBER: 60/249,211
;  PRIOR FILING DATE: 2000-11-17
;  PRIOR APPLICATION NUMBER: 60/249,215
;  PRIOR FILING DATE: 2000-11-17
;  PRIOR APPLICATION NUMBER: 60/249,264
;  PRIOR FILING DATE: 2000-11-17
;  PRIOR APPLICATION NUMBER: 60/249,214
;  PRIOR FILING DATE: 2000-11-17
;  PRIOR APPLICATION NUMBER: 60/249,297
;  PRIOR FILING DATE: 2000-11-17
;  PRIOR APPLICATION NUMBER: 60/232,400
;  PRIOR FILING DATE: 2000-09-14
;  PRIOR APPLICATION NUMBER: 60/231,242
;  PRIOR FILING DATE: 2000-09-08
;  PRIOR APPLICATION NUMBER: 60/232,081
;  PRIOR FILING DATE: 2000-09-08
;  PRIOR APPLICATION NUMBER: 60/232,080
;  PRIOR FILING DATE: 2000-09-08
;  PRIOR APPLICATION NUMBER: 60/231,414
;  PRIOR FILING DATE: 2000-09-08
;  PRIOR APPLICATION NUMBER: 60/231,244
;  PRIOR FILING DATE: 2000-09-08
;  PRIOR APPLICATION NUMBER: 60/233,064
;  PRIOR FILING DATE: 2000-09-14
;  PRIOR APPLICATION NUMBER: 60/233,063
;  PRIOR FILING DATE: 2000-09-14
;  PRIOR APPLICATION NUMBER: 60/232,397
```

```
;  PRIOR FILING DATE: 2000-09-14
;  PRIOR APPLICATION NUMBER: 60/232,399
;  PRIOR FILING DATE: 2000-09-14
;  PRIOR APPLICATION NUMBER: 60/232,401
;  PRIOR FILING DATE: 2000-09-14
;  PRIOR APPLICATION NUMBER: 60/241,808
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/241,826
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/241,786
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/241,221
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/246,475
;  PRIOR FILING DATE: 2000-11-08
;  PRIOR APPLICATION NUMBER: 60/231,243
;  PRIOR FILING DATE: 2000-09-08
;  PRIOR APPLICATION NUMBER: 60/233,065
;  PRIOR FILING DATE: 2000-09-14
;  PRIOR APPLICATION NUMBER: 60/232,398

  Query Match             65.5%;  Score 38;  DB 5;  Length 114;
  Best Local Similarity   70.0%;  Pred. No. 89;
  Matches    7;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy        1 HLMPFPGPLL 10
            ||  |||| |
Db       74 HLASFPGPSL 83


RESULT 11
US-09-950-083-3140
; Sequence 3140, Application US/09950083
; GENERAL INFORMATION:
;  APPLICANT: Rosen, et. al
;  TITLE OF INVENTION: Human Secreted Proteins
;  FILE REFERENCE: PS805
;  CURRENT APPLICATION NUMBER: US/09/950,083
;  CURRENT FILING DATE:  2001-09-12
;  PRIOR APPLICATION NUMBER: 60/278,650
;  PRIOR FILING DATE: 2001-03-27
;  PRIOR APPLICATION NUMBER: 09/833,245
;  PRIOR FILING DATE: 2001-04-12
;  PRIOR APPLICATION NUMBER: PCT/US01/11988
;  PRIOR FILING DATE: 2001-04-12
;  PRIOR APPLICATION NUMBER: PCT/US00/06043
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06012
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06058
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06044
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06059
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06042
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06014
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06013
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06049
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06057
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06824
;  PRIOR FILING DATE: 2000-03-16
;  PRIOR APPLICATION NUMBER: PCT/US00/06765
;  PRIOR FILING DATE: 2000-03-16
;  PRIOR APPLICATION NUMBER: PCT/US00/06792
;  PRIOR FILING DATE: 2000-03-16
;  PRIOR APPLICATION NUMBER: PCT/US00/06830
;  PRIOR FILING DATE: 2000-03-16
```

```
Qy      1 HLMPFPGPLL 10
          ||  |||| |
Db     20 HLASFPGPSL 29


RESULT   9
US-09-758-462-1327
; Sequence 1327, Application US/09758462
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PM007
;  CURRENT APPLICATION NUMBER: US/09/758,462
;  CURRENT FILING DATE:  2001-01-11
;  PRIOR APPLICATION NUMBER: 60/179,065
;  PRIOR FILING DATE: 2000-01-31
;  PRIOR APPLICATION NUMBER: 60/180,628
;  PRIOR FILING DATE: 2000-02-04
;  NUMBER OF SEQ ID NOS: 1734
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1327
;   LENGTH: 83
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-758-462-1327


  Query Match             65.5%;  Score 38;  DB 5;  Length 83;
  Best Local Similarity   85.7%;  Pred. No. 67;
  Matches    6;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      4 PFPGPLL 10
          ||||||:
Db     26 PFPGPLV 32


RESULT  10
US-09-764-905-12711
; Sequence 12711, Application US/09764905
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PC004
;  CURRENT APPLICATION NUMBER: US/09/764,905
;  CURRENT FILING DATE:  2001-01-17
;  PRIOR APPLICATION NUMBER: 60/179,065
;  PRIOR FILING DATE: 2000-01-31
;  PRIOR APPLICATION NUMBER: 60/180,628
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: 60/214,886
;  PRIOR FILING DATE: 2000-06-28
;  PRIOR APPLICATION NUMBER: 60/217,487
;  PRIOR FILING DATE: 2000-07-11
;  PRIOR APPLICATION NUMBER: 60/225,758
;  PRIOR FILING DATE: 2000-08-14
;  PRIOR APPLICATION NUMBER: 60/220,963
;  PRIOR FILING DATE: 2000-07-26
;  PRIOR APPLICATION NUMBER: 60/217,496
;  PRIOR FILING DATE: 2000-07-11
;  PRIOR APPLICATION NUMBER: 60/225,447
;  PRIOR FILING DATE: 2000-08-14
;  PRIOR APPLICATION NUMBER: 60/218,290
;  PRIOR FILING DATE: 2000-07-14
;  PRIOR APPLICATION NUMBER: 60/225,757
;  PRIOR FILING DATE: 2000-08-14
;  PRIOR APPLICATION NUMBER: 60/226,868
;  PRIOR FILING DATE: 2000-08-22
;  PRIOR APPLICATION NUMBER: 60/216,647
;  PRIOR FILING DATE: 2000-07-07
;  PRIOR APPLICATION NUMBER: 60/225,267
;  PRIOR FILING DATE: 2000-08-14
;  PRIOR APPLICATION NUMBER: 60/216,880
;  PRIOR FILING DATE: 2000-07-07
;  PRIOR APPLICATION NUMBER: 60/225,270
;  PRIOR FILING DATE: 2000-08-14
;  PRIOR APPLICATION NUMBER: 60/251,869
;  PRIOR FILING DATE: 2000-12-08
;  PRIOR APPLICATION NUMBER: 60/235,834
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: 60/234,274
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: 60/234,223
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: 60/228,924
;  PRIOR FILING DATE: 2000-08-30
;  PRIOR APPLICATION NUMBER: 60/224,518
;  PRIOR FILING DATE: 2000-08-14
;  PRIOR APPLICATION NUMBER: 60/236,369
;  PRIOR FILING DATE: 2000-09-29
;  PRIOR APPLICATION NUMBER: 60/224,519
;  PRIOR FILING DATE: 2000-08-14
;  PRIOR APPLICATION NUMBER: 60/220,964
;  PRIOR FILING DATE: 2000-07-26
;  PRIOR APPLICATION NUMBER: 60/241,809
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/249,299
;  PRIOR FILING DATE: 2000-11-17
;  PRIOR APPLICATION NUMBER: 60/236,327
;  PRIOR FILING DATE: 2000-09-29
;  PRIOR APPLICATION NUMBER: 60/241,785
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/244,617
;  PRIOR FILING DATE: 2000-11-01
;  PRIOR APPLICATION NUMBER: 60/225,268
;  PRIOR FILING DATE: 2000-08-14
;  PRIOR APPLICATION NUMBER: 60/236,368
;  PRIOR FILING DATE: 2000-09-29
;  PRIOR APPLICATION NUMBER: 60/251,856
;  PRIOR FILING DATE: 2000-12-08
;  PRIOR APPLICATION NUMBER: 60/251,868
;  PRIOR FILING DATE: 2000-12-08
;  PRIOR APPLICATION NUMBER: 60/229,344
;  PRIOR FILING DATE: 2000-09-01
;  PRIOR APPLICATION NUMBER: 60/234,997
;  PRIOR FILING DATE: 2000-09-25
;  PRIOR APPLICATION NUMBER: 60/229,343
;  PRIOR FILING DATE: 2000-09-01
;  PRIOR APPLICATION NUMBER: 60/229,345
;  PRIOR FILING DATE: 2000-09-01
;  PRIOR APPLICATION NUMBER: 60/229,287
;  PRIOR FILING DATE: 2000-09-01
;  PRIOR APPLICATION NUMBER: 60/229,513
;  PRIOR FILING DATE: 2000-09-05
;  PRIOR APPLICATION NUMBER: 60/231,413
;  PRIOR FILING DATE: 2000-09-08
;  PRIOR APPLICATION NUMBER: 60/229,509
;  PRIOR FILING DATE: 2000-09-05
;  PRIOR APPLICATION NUMBER: 60/236,367
;  PRIOR FILING DATE: 2000-09-29
;  PRIOR APPLICATION NUMBER: 60/237,039
;  PRIOR FILING DATE: 2000-10-02
;  PRIOR APPLICATION NUMBER: 60/237,038
;  PRIOR FILING DATE: 2000-10-02
;  PRIOR APPLICATION NUMBER: 60/236,370
;  PRIOR FILING DATE: 2000-09-29
;  PRIOR APPLICATION NUMBER: 60/236,802
;  PRIOR FILING DATE: 2000-10-02
;  PRIOR APPLICATION NUMBER: 60/237,037
;  PRIOR FILING DATE: 2000-10-02
;  PRIOR APPLICATION NUMBER: 60/237,040
;  PRIOR FILING DATE: 2000-10-02
;  PRIOR APPLICATION NUMBER: 60/240,960
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/239,935
;  PRIOR FILING DATE: 2000-10-13
```

```
;   OTHER INFORMATION: Xaa is any amino acid
US-09-617-681A-711


  Query Match             67.2%;  Score 39;  DB 5;  Length 61;
  Best Local Similarity   85.7%;  Pred. No. 36;
  Matches    6;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 MPFPGPL 9
              :||||||
Db          3 LPFPGPL 9


RESULT 6
US-09-757-028-2555
; Sequence 2555, Application US/09757028
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PM001
;  CURRENT APPLICATION NUMBER: US/09/757,028
;  CURRENT FILING DATE:  2001-01-09
;  PRIOR APPLICATION NUMBER: 60/179,065
;  PRIOR FILING DATE: 2000-01-31
;  PRIOR APPLICATION NUMBER: 60/180,628
;  PRIOR FILING DATE: 2000-02-04
;  NUMBER OF SEQ ID NOS: 2660
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2555
;   LENGTH: 78
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-757-028-2555


  Query Match             67.2%;  Score 39;  DB 5;  Length 78;
  Best Local Similarity   70.0%;  Pred. No. 45;
  Matches    7;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 HLMPFPGPLL 10
              ||  :|||||
Db         13 HLSTWPGPLL 22


RESULT 7
US-09-864-761-48232
; Sequence 48232, Application US/09864761
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
```

```
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 48232
;   LENGTH: 104
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AC002504.1
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.42
;   OTHER INFORMATION: EST_HUMAN HIT: AW028575.1, EVALUE 6.00e-19
;   OTHER INFORMATION: SWISSPROT HIT: O35085, EVALUE 2.00e-29
US-09-864-761-48232


  Query Match             67.2%;  Score 39;  DB 5;  Length 104;
  Best Local Similarity   85.7%;  Pred. No. 58;
  Matches    6;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 MPFPGPL 9
              :||||||
Db         20 LPFPGPL 26


RESULT 8
US-09-834-366-19427
; Sequence 19427, Application US/09834366
; GENERAL INFORMATION:
;  APPLICANT: Bejanin, Stephane
;  APPLICANT:  Tanaka, Hiroaki
;  APPLICANT:  Dumas Milne Edwards, Jean Baptiste
;  APPLICANT:  Jobert, Severin
;  APPLICANT:  Giordano, Jean-Yves
;  TITLE OF INVENTION: ESTs and Encoded Human Proteins.
;  FILE REFERENCE: 81.US2.REG
;  CURRENT APPLICATION NUMBER: US/09/834,366
;  CURRENT FILING DATE:  2001-04-13
;  PRIOR APPLICATION NUMBER: US 60/197,873
;  PRIOR FILING DATE: 2000-04-18
;  NUMBER OF SEQ ID NOS: 52153
;  SOFTWARE: Patent.pm
; SEQ ID NO 19427
;   LENGTH: 50
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SIGNAL
;   LOCATION: -37..-1
US-09-834-366-19427


  Query Match             65.5%;  Score 38;  DB 5;  Length 50;
  Best Local Similarity   70.0%;  Pred. No. 42;
  Matches    7;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;
```

```
;  PRIOR FILING DATE: 2000-11-08
;  PRIOR APPLICATION NUMBER: 60/231,243
;  PRIOR FILING DATE: 2000-09-08
;  PRIOR APPLICATION NUMBER: 60/233,065
;  PRIOR FILING DATE: 2000-09-14
;  PRIOR APPLICATION NUMBER: 60/232,398

  Query Match             75.9%;  Score 44;  DB 5;  Length 98;
  Best Local Similarity   60.0%;  Pred. No. 10;
  Matches    6;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 HLMPFPGPLL 10
           |::||| |:|
Db      42 HVLPFPSPML 51


RESULT   2
PCT-US01-08631-60550
; Sequence 60550, Application PC/TUS0108631
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc
;  TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
;  FILE REFERENCE: 21272-049
;  CURRENT APPLICATION NUMBER: PCT/US01/08631
;  CURRENT FILING DATE:  2001-03-30
;  PRIOR APPLICATION NUMBER: 09/540,217
;  PRIOR FILING DATE: 2000-03-31
;  PRIOR APPLICATION NUMBER: 09/649,167
;  PRIOR FILING DATE: 2000-08-23
;  NUMBER OF SEQ ID NOS: 60736
;  SOFTWARE: Custom
; SEQ ID NO 60550
;   LENGTH: 217
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: DOMAIN
;   LOCATION: (126)..(142)
;   OTHER INFORMATION: Purine/pyrimidine phosphoribosyl transferases proteins domain
;   OTHER INFORMATION: identified by eMATRIX, accession number BL00103B, p-value=1.947e-
;   OTHER INFORMATION: 14, raw score of 9.64
;   NAME/KEY: DOMAIN
;   LOCATION: (34)..(188)
;   OTHER INFORMATION: Phosphoribosyl transferase domain identified by PFam,
;   OTHER INFORMATION: accession name Pribosyltran, E-value=3.3e-43, PFam score of 157.0
PCT-US01-08631-60550


  Query Match             72.4%;  Score 42;  DB 1;  Length 217;
  Best Local Similarity   77.8%;  Pred. No. 41;
  Matches    7;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 LMPFPGPLL 10
           | ||||||:
Db     184 LCPFPGPLI 192


RESULT   3
US-09-803-110-9641
; Sequence 9641, Application US/09803110
; GENERAL INFORMATION:
;  APPLICANT: Hinkle, Gregory J.
;  APPLICANT:  Slater, Steven C.
;  TITLE OF INVENTION: Agrobacterium tumefaciens Genome Sequences and Uses Thereof
;  FILE REFERENCE: 38-10(15490)D
;  CURRENT APPLICATION NUMBER: US/09/803,110
;  CURRENT FILING DATE:  2001-03-12
;  PRIOR APPLICATION NUMBER: US 09/739,449
;  PRIOR FILING DATE: 2000-12-19
;  PRIOR APPLICATION NUMBER: US 09/514,000
;  PRIOR FILING DATE: 2000-02-23
;  PRIOR APPLICATION NUMBER: US 60/168,139
;  PRIOR FILING DATE: 1999-12-01
```

```
;  NUMBER OF SEQ ID NOS: 13351
; SEQ ID NO 9641
;   LENGTH: 534
;   TYPE: PRT
;   ORGANISM: Agrobacterium tumefaciens
US-09-803-110-9641


  Query Match             70.7%;  Score 41;  DB 5;  Length 534;
  Best Local Similarity   75.0%;  Pred. No. 1.3e+02;
  Matches    6;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 HLMPFPGP 8
           ||||:| |
Db     260 HLMPYPAP 267


RESULT   4
PCT-US01-08631-44470
; Sequence 44470, Application PC/TUS0108631
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc
;  TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
;  FILE REFERENCE: 21272-049
;  CURRENT APPLICATION NUMBER: PCT/US01/08631
;  CURRENT FILING DATE:  2001-03-30
;  PRIOR APPLICATION NUMBER: 09/540,217
;  PRIOR FILING DATE: 2000-03-31
;  PRIOR APPLICATION NUMBER: 09/649,167
;  PRIOR FILING DATE: 2000-08-23
;  NUMBER OF SEQ ID NOS: 60736
;  SOFTWARE: Custom
; SEQ ID NO 44470
;   LENGTH: 547
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(547)
;   OTHER INFORMATION: Xaa = X or * as defined in Table 2
PCT-US01-08631-44470


  Query Match             70.7%;  Score 41;  DB 1;  Length 547;
  Best Local Similarity   77.8%;  Pred. No. 1.3e+02;
  Matches    7;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 HLMPFPGPL 9
           ||:| ||||
Db     405 HLVPGPGPL 413


RESULT   5
US-09-617-681A-711
; Sequence 711, Application US/09617681A
; GENERAL INFORMATION:
;  APPLICANT: N. ALEXANDROV et al.
;  TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptid
;  TITLE OF INVENTION:  Thereby
;  FILE REFERENCE: 2750-1064P
;  CURRENT APPLICATION NUMBER: US/09/617,681A
;  CURRENT FILING DATE:  2000-07-19
;  NUMBER OF SEQ ID NOS: 9311
; SEQ ID NO 711
;   LENGTH: 61
;   TYPE: PRT
;   ORGANISM: Arabidopsis thaliana
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 1..61
;   OTHER INFORMATION: Ceres Seq. ID 1426763
;   NAME/KEY: misc_feature
;   LOCATION: 1..61
```

```
;  PRIOR FILING DATE: 2000-08-14
;  PRIOR APPLICATION NUMBER: 60/236,369
;  PRIOR FILING DATE: 2000-09-29
;  PRIOR APPLICATION NUMBER: 60/224,519
;  PRIOR FILING DATE: 2000-08-14
;  PRIOR APPLICATION NUMBER: 60/220,964
;  PRIOR FILING DATE: 2000-07-26
;  PRIOR APPLICATION NUMBER: 60/241,809
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/249,299
;  PRIOR FILING DATE: 2000-11-17
;  PRIOR APPLICATION NUMBER: 60/236,327
;  PRIOR FILING DATE: 2000-09-29
;  PRIOR APPLICATION NUMBER: 60/241,785
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/244,617
;  PRIOR FILING DATE: 2000-11-01
;  PRIOR APPLICATION NUMBER: 60/225,268
;  PRIOR FILING DATE: 2000-08-14
;  PRIOR APPLICATION NUMBER: 60/236,368
;  PRIOR FILING DATE: 2000-09-29
;  PRIOR APPLICATION NUMBER: 60/251,856
;  PRIOR FILING DATE: 2000-12-08
;  PRIOR APPLICATION NUMBER: 60/251,868
;  PRIOR FILING DATE: 2000-12-08
;  PRIOR APPLICATION NUMBER: 60/229,344
;  PRIOR FILING DATE: 2000-09-01
;  PRIOR APPLICATION NUMBER: 60/234,997
;  PRIOR FILING DATE: 2000-09-25
;  PRIOR APPLICATION NUMBER: 60/229,343
;  PRIOR FILING DATE: 2000-09-01
;  PRIOR APPLICATION NUMBER: 60/229,345
;  PRIOR FILING DATE: 2000-09-01
;  PRIOR APPLICATION NUMBER: 60/229,287
;  PRIOR FILING DATE: 2000-09-01
;  PRIOR APPLICATION NUMBER: 60/229,513
;  PRIOR FILING DATE: 2000-09-05
;  PRIOR APPLICATION NUMBER: 60/231,413
;  PRIOR FILING DATE: 2000-09-08
;  PRIOR APPLICATION NUMBER: 60/229,509
;  PRIOR FILING DATE: 2000-09-05
;  PRIOR APPLICATION NUMBER: 60/236,367
;  PRIOR FILING DATE: 2000-09-29
;  PRIOR APPLICATION NUMBER: 60/237,039
;  PRIOR FILING DATE: 2000-10-02
;  PRIOR APPLICATION NUMBER: 60/237,038
;  PRIOR FILING DATE: 2000-10-02
;  PRIOR APPLICATION NUMBER: 60/236,370
;  PRIOR FILING DATE: 2000-09-29
;  PRIOR APPLICATION NUMBER: 60/236,802
;  PRIOR FILING DATE: 2000-10-02
;  PRIOR APPLICATION NUMBER: 60/237,037
;  PRIOR FILING DATE: 2000-10-02
;  PRIOR APPLICATION NUMBER: 60/237,040
;  PRIOR FILING DATE: 2000-10-02
;  PRIOR APPLICATION NUMBER: 60/240,960
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/239,935
;  PRIOR FILING DATE: 2000-10-13
;  PRIOR APPLICATION NUMBER: 60/239,937
;  PRIOR FILING DATE: 2000-10-13
;  PRIOR APPLICATION NUMBER: 60/241,787
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/246,474
;  PRIOR FILING DATE: 2000-11-08
;  PRIOR APPLICATION NUMBER: 60/246,532
;  PRIOR FILING DATE: 2000-11-08
;  PRIOR APPLICATION NUMBER: 60/249,216
;  PRIOR FILING DATE: 2000-11-17
;  PRIOR APPLICATION NUMBER: 60/249,210
;  PRIOR FILING DATE: 2000-11-17
;  PRIOR APPLICATION NUMBER: 60/226,681
;  PRIOR FILING DATE: 2000-08-22
```

```
;  PRIOR APPLICATION NUMBER: 60/225,759
;  PRIOR FILING DATE: 2000-08-14
;  PRIOR APPLICATION NUMBER: 60/225,213
;  PRIOR FILING DATE: 2000-08-14
;  PRIOR APPLICATION NUMBER: 60/227,182
;  PRIOR FILING DATE: 2000-08-22
;  PRIOR APPLICATION NUMBER: 60/225,214
;  PRIOR FILING DATE: 2000-08-14
;  PRIOR APPLICATION NUMBER: 60/235,836
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: 60/230,438
;  PRIOR FILING DATE: 2000-09-06
;  PRIOR APPLICATION NUMBER: 60/215,135
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: 60/225,266
;  PRIOR FILING DATE: 2000-08-14
;  PRIOR APPLICATION NUMBER: 60/249,218
;  PRIOR FILING DATE: 2000-11-17
;  PRIOR APPLICATION NUMBER: 60/249,208
;  PRIOR FILING DATE: 2000-11-17
;  PRIOR APPLICATION NUMBER: 60/249,213
;  PRIOR FILING DATE: 2000-11-17
;  PRIOR APPLICATION NUMBER: 60/249,212
;  PRIOR FILING DATE: 2000-11-17
;  PRIOR APPLICATION NUMBER: 60/249,207
;  PRIOR FILING DATE: 2000-11-17
;  PRIOR APPLICATION NUMBER: 60/249,245
;  PRIOR FILING DATE: 2000-11-17
;  PRIOR APPLICATION NUMBER: 60/249,244
;  PRIOR FILING DATE: 2000-11-17
;  PRIOR APPLICATION NUMBER: 60/249,217
;  PRIOR FILING DATE: 2000-11-17
;  PRIOR APPLICATION NUMBER: 60/249,211
;  PRIOR FILING DATE: 2000-11-17
;  PRIOR APPLICATION NUMBER: 60/249,215
;  PRIOR FILING DATE: 2000-11-17
;  PRIOR APPLICATION NUMBER: 60/249,264
;  PRIOR FILING DATE: 2000-11-17
;  PRIOR APPLICATION NUMBER: 60/249,214
;  PRIOR FILING DATE: 2000-11-17
;  PRIOR APPLICATION NUMBER: 60/249,297
;  PRIOR FILING DATE: 2000-11-17
;  PRIOR APPLICATION NUMBER: 60/232,400
;  PRIOR FILING DATE: 2000-09-14
;  PRIOR APPLICATION NUMBER: 60/231,242
;  PRIOR FILING DATE: 2000-09-08
;  PRIOR APPLICATION NUMBER: 60/232,081
;  PRIOR FILING DATE: 2000-09-08
;  PRIOR APPLICATION NUMBER: 60/232,080
;  PRIOR FILING DATE: 2000-09-08
;  PRIOR APPLICATION NUMBER: 60/231,414
;  PRIOR FILING DATE: 2000-09-08
;  PRIOR APPLICATION NUMBER: 60/231,244
;  PRIOR FILING DATE: 2000-09-08
;  PRIOR APPLICATION NUMBER: 60/233,064
;  PRIOR FILING DATE: 2000-09-14
;  PRIOR APPLICATION NUMBER: 60/233,063
;  PRIOR FILING DATE: 2000-09-14
;  PRIOR APPLICATION NUMBER: 60/232,397
;  PRIOR FILING DATE: 2000-09-14
;  PRIOR APPLICATION NUMBER: 60/232,399
;  PRIOR FILING DATE: 2000-09-14
;  PRIOR APPLICATION NUMBER: 60/232,401
;  PRIOR FILING DATE: 2000-09-14
;  PRIOR APPLICATION NUMBER: 60/241,808
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/241,826
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/241,786
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/241,221
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/246,475
```

```
                        GenCore version 4.5
              Copyright (c) 1993 - 2000  Compugen Ltd.


OM protein - protein search, using sw model

Run on:         October 23, 2001, 13:22:51 ; Search time 39.01 Seconds
                                             (without alignments)
                                             13.209 Million cell updates/sec

Title:          US-09-625-963-3
Perfect score:  58
Sequence:       1 HLMPFPGPLL 10

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched:       339663 seqs, 51527020 residues

Total number of hits satisfying chosen parameters:      339663

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       Pending_Patents_AA_New:*
               1:  /cgn2_6/ptodata/2/paa/PCT_NEW_COMB.pep:*
               2:  /cgn2_6/ptodata/2/paa/US06_NEW_COMB.pep:*
               3:  /cgn2_6/ptodata/2/paa/US07_NEW_COMB.pep:*
               4:  /cgn2_6/ptodata/2/paa/US08_NEW_COMB.pep:*
               5:  /cgn2_6/ptodata/2/paa/US09_NEW_COMB.pep:*
               6:  /cgn2_6/ptodata/2/paa/US60_NEW_COMB.pep:*

     Pred. No. is the number of results predicted by chance to have a
     score greater than or equal to the score of the result being printed,
     and is derived by analysis of the total score distribution.

                              SUMMARIES
                 %
Result          Query
  No.   Score   Match Length DB  ID                      Description
-------------------------------------------------------------------------
    1      44    75.9     98  5  US-09-764-905-15768       Sequence 15768, A
    2      42    72.4    217  1  PCT-US01-08631-60550      Sequence 60550, A
    3      41    70.7    534  5  US-09-803-110-9641        Sequence 9641, Ap
    4      41    70.7    547  1  PCT-US01-08631-44470      Sequence 44470, A
    5      39    67.2     61  5  US-09-617-681A-711        Sequence 711, App
    6      39    67.2     78  5  US-09-757-028-2555        Sequence 2555, Ap
    7      39    67.2    104  5  US-09-864-761-48232       Sequence 48232, A
    8      38    65.5     50  5  US-09-834-366-19427       Sequence 19427, A
    9      38    65.5     83  5  US-09-758-462-1327        Sequence 1327, Ap
   10      38    65.5    114  5  US-09-764-905-12711       Sequence 12711, A
   11      38    65.5    117  5  US-09-950-083-3140        Sequence 3140, Ap
   12      38    65.5    161  5  US-09-617-682A-3093       Sequence 3093, Ap
   13      38    65.5    450  6  US-60-312-544-6518        Sequence 6518, Ap
   14      37    63.8     47  1  PCT-US01-08656-9992       Sequence 9992, Ap
   15      37    63.8     88  5  US-09-758-462-903         Sequence 903, App
   16      37    63.8    784  6  US-60-311-261-2800        Sequence 2800, Ap
   17      37    63.8   1295  1  PCT-US01-14827-15531      Sequence 15531, A
   18      37    63.8   1646  1  PCT-US01-14827-15530      Sequence 15530, A
   19      37    63.8   1911  5  US-09-854-856-64          Sequence 64, Appl
   20      37    63.8   1939  5  US-09-854-856-48          Sequence 48, Appl
   21      37    63.8   1971  5  US-09-854-856-32          Sequence 32, Appl
   22      37    63.8   1999  5  US-09-854-856-16          Sequence 16, Appl
   23      37    63.8   2004  5  US-09-854-856-58          Sequence 58, Appl
   24      37    63.8   2032  5  US-09-854-856-42          Sequence 42, Appl
   25      37    63.8   2048  5  US-09-854-856-62          Sequence 62, Appl
   26      37    63.8   2064  5  US-09-854-856-26          Sequence 26, Appl
   27      37    63.8   2076  5  US-09-854-856-46          Sequence 46, Appl
```

```
   28      37    63.8   2092  5  US-09-854-856-10          Sequence 10, Appl
   29      37    63.8   2108  5  US-09-854-856-30          Sequence 30, Appl
   30      37    63.8   2136  5  US-09-854-856-14          Sequence 14, Appl
   31      37    63.8   2141  5  US-09-854-856-56          Sequence 56, Appl
   32      37    63.8   2157  5  US-09-854-856-52          Sequence 52, Appl
   33      37    63.8   2169  5  US-09-854-856-40          Sequence 40, Appl
   34      37    63.8   2185  5  US-09-854-856-36          Sequence 36, Appl
   35      37    63.8   2201  5  US-09-854-856-24          Sequence 24, Appl
   36      37    63.8   2217  5  US-09-854-856-20          Sequence 20, Appl
   37      37    63.8   2229  5  US-09-854-856-8           Sequence 8, Appli
   38      37    63.8   2245  5  US-09-854-856-4           Sequence 4, Appli
   39      37    63.8   2294  5  US-09-854-856-50          Sequence 50, Appl
   40      37    63.8   2322  5  US-09-854-856-34          Sequence 34, Appl
   41      37    63.8   2354  5  US-09-854-856-18          Sequence 18, Appl
   42      37    63.8   2382  5  US-09-854-856-2           Sequence 2, Appli
   43    36.5    62.9    644  1  PCT-US01-08656-7313       Sequence 7313, Ap
   44      36    62.1     32  5  US-09-617-682A-15882      Sequence 15882, A
   45      36    62.1     52  5  US-09-617-682A-11585      Sequence 11585, A


                              ALIGNMENTS


RESULT   1
US-09-764-905-15768
; Sequence 15768, Application US/09764905
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PC004
;  CURRENT APPLICATION NUMBER: US/09/764,905
;  CURRENT FILING DATE:  2001-01-17
;  PRIOR APPLICATION NUMBER: 60/179,065
;  PRIOR FILING DATE: 2000-01-31
;  PRIOR APPLICATION NUMBER: 60/180,628
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: 60/214,886
;  PRIOR FILING DATE: 2000-06-28
;  PRIOR APPLICATION NUMBER: 60/217,487
;  PRIOR FILING DATE: 2000-07-11
;  PRIOR APPLICATION NUMBER: 60/225,758
;  PRIOR FILING DATE: 2000-08-14
;  PRIOR APPLICATION NUMBER: 60/220,963
;  PRIOR FILING DATE: 2000-07-26
;  PRIOR APPLICATION NUMBER: 60/217,496
;  PRIOR FILING DATE: 2000-07-11
;  PRIOR APPLICATION NUMBER: 60/225,447
;  PRIOR FILING DATE: 2000-08-14
;  PRIOR APPLICATION NUMBER: 60/218,290
;  PRIOR FILING DATE: 2000-07-14
;  PRIOR APPLICATION NUMBER: 60/225,757
;  PRIOR FILING DATE: 2000-08-14
;  PRIOR APPLICATION NUMBER: 60/226,868
;  PRIOR FILING DATE: 2000-08-22
;  PRIOR APPLICATION NUMBER: 60/216,647
;  PRIOR FILING DATE: 2000-07-07
;  PRIOR APPLICATION NUMBER: 60/225,267
;  PRIOR FILING DATE: 2000-08-14
;  PRIOR APPLICATION NUMBER: 60/216,880
;  PRIOR FILING DATE: 2000-07-07
;  PRIOR APPLICATION NUMBER: 60/225,270
;  PRIOR FILING DATE: 2000-08-14
;  PRIOR APPLICATION NUMBER: 60/251,869
;  PRIOR FILING DATE: 2000-12-08
;  PRIOR APPLICATION NUMBER: 60/235,834
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: 60/234,274
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: 60/234,223
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: 60/228,924
;  PRIOR FILING DATE: 2000-08-30
;  PRIOR APPLICATION NUMBER: 60/224,518
```

```
; SEQ ID NO 40482
;   LENGTH: 103
;   TYPE: PRT
;   ORGANISM: A.fumigatus
US-09-417-507-40482


  Query Match             67.2%;  Score 39;  DB 18;  Length 103;
  Best Local Similarity   87.5%;  Pred. No. 3.2e+02;
  Matches    7;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        3 MPFPGPLL 10
            |||||| |
Db       42 MPFPGPPL 49


Search completed: October 23, 2001, 13:27:00
Job time: 413 sec
```

```
  Query Match             69.0%;  Score 40;  DB 23;  Length 345;
  Best Local Similarity   60.0%;  Pred. No. 7.4e+02;
  Matches    6;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 HLMPFPGPLL 10
          |:|| | ||:
Db     39 HIMPAPSPLI 48


RESULT  11
US-60-191-637-7407
; Sequence 7407, Application US/60191637
; GENERAL INFORMATION:
;  APPLICANT: Venter, J. Craig
;  TITLE OF INVENTION: NUCLEIC ACID DETECTION KITS COMPRISING
;  TITLE OF INVENTION: GENE SEQUENCES EXPRESSED FROM THE DROSOPHILA GENOME, AND
;  TITLE OF INVENTION: USES THEREOF
;  FILE REFERENCE: CL000392
;  CURRENT APPLICATION NUMBER: US/60/191,637
;  CURRENT FILING DATE: 2000-03-23
;  NUMBER OF SEQ ID NOS: 42660
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7407
;   LENGTH: 345
;   TYPE: PRT
;   ORGANISM: DROSOPHILA
US-60-191-637-7407


  Query Match             69.0%;  Score 40;  DB 23;  Length 345;
  Best Local Similarity   60.0%;  Pred. No. 7.4e+02;
  Matches    6;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 HLMPFPGPLL 10
          |:|| | ||:
Db     39 HIMPAPSPLI 48


RESULT  12
US-09-134-000-4944
; Sequence 4944, Application US/09134000A
; GENERAL INFORMATION:
;  APPLICANT: Lynn Doucette-Stamm et al
;  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ENTEROCOCCUS
;  TITLE OF INVENTION: FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE: GTC-005
;  CURRENT APPLICATION NUMBER: US/09/134,000A
;  CURRENT FILING DATE: 1998-08-13
;  NUMBER OF SEQ ID NOS: 6810
; SEQ ID NO 4944
;   LENGTH: 449
;   TYPE: PRT
;   ORGANISM: Enterococcus faecalis
US-09-134-000-4944


  Query Match             69.0%;  Score 40;  DB 15;  Length 449;
  Best Local Similarity   70.0%;  Pred. No. 9.6e+02;
  Matches    7;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 HLMPFPGPLL 10
          ||  ||||:|
Db    298 HLTGFPGPVL 307


RESULT  13
US-09-527-522-5
; Sequence 5, Application US/09527522
; GENERAL INFORMATION:
;  APPLICANT:  Takashima, Yoshiki
;  APPLICANT:  Mitsuda, Satoshi
;  TITLE OF INVENTION:  Protein capable of catalyzing transamination
;  TITLE OF INVENTION:  stereoselectively, gene encoding said protein
;  TITLE OF INVENTION:  and use thereof
;  FILE REFERENCE:  Q58251
;  CURRENT APPLICATION NUMBER: US/09/527,522
;  CURRENT FILING DATE:  2000-03-16
;  PRIOR APPLICATION NUMBER:  JP 11/075511
;  PRIOR FILING DATE:  1999-03-19
;  PRIOR APPLICATION NUMBER:  JP 11/088634
;  PRIOR FILING DATE:  1999-03-30
;  NUMBER OF SEQ ID NOS:  12
;  SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 5
;   LENGTH: 19
;   TYPE: PRT
;   ORGANISM: Mycobacterium aurum
;   FEATURE:
;   OTHER INFORMATION: SC-S423
US-09-527-522-5


  Query Match             67.2%;  Score 39;  DB 19;  Length 19;
  Best Local Similarity   70.0%;  Pred. No. 60;
  Matches    7;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 HLMPFPGPLL 10
          |||  ||||:
Db      5 HLMDEPGPLI 14


RESULT  14
US-60-185-361-923
; Sequence 923, Application US/60185361
; GENERAL INFORMATION:
;  APPLICANT: Bonazzi, Vivien
;  TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
;  TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS, AND
;  TITLE OF INVENTION: USES THEREOF
;  FILE REFERENCE: CL000291
;  CURRENT APPLICATION NUMBER: US/60/185,361
;  CURRENT FILING DATE: 2000-02-28
;  NUMBER OF SEQ ID NOS: 968
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 923
;   LENGTH: 77
;   TYPE: PRT
;   ORGANISM: HUMAN
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: (1)...(77)
;   OTHER INFORMATION: Xaa = Any Amino Acid
US-60-185-361-923


  Query Match             67.2%;  Score 39;  DB 23;  Length 77;
  Best Local Similarity   77.8%;  Pred. No. 2.4e+02;
  Matches    7;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      2 LMPFPGPLL 10
          |:|| ||||
Db     33 LLPFQGPLL 41


RESULT  15
US-09-417-507-40482
; Sequence 40482, Application US/09417507
; GENERAL INFORMATION:
;  APPLICANT: KEITH G. WEINSTOCK ET AL.
;  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ASPERGILLUS
;  TITLE OF INVENTION: FUMIGATUS FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE: PATH99-10
;  CURRENT APPLICATION NUMBER: US/09/417,507
;  CURRENT FILING DATE: 1999-10-14
;  NUMBER OF SEQ ID NOS: 44312
```

```
Qy      3 MPFPGPLL 10
          :||||||:
Db     54 LPFPGPLI 61


RESULT   6
US-09-739-449-9641
; Sequence 9641, Application US/09739449
; GENERAL INFORMATION:
;  APPLICANT: Hinkle, Gregory J.
;  APPLICANT:  Slater, Steven C.
;  TITLE OF INVENTION: Agrobacterium tumefaciens Genome Sequences and Uses Thereof
;  FILE REFERENCE: 38-10(15490)C
;  CURRENT APPLICATION NUMBER: US/09/739,449
;  CURRENT FILING DATE:  2000-12-19
;  PRIOR APPLICATION NUMBER: US 09/514,000
;  PRIOR FILING DATE: 2000-02-23
;  NUMBER OF SEQ ID NOS: 13351
; SEQ ID NO 9641
;   LENGTH: 534
;   TYPE: PRT
;   ORGANISM: Agrobacterium tumefaciens
US-09-739-449-9641


  Query Match             70.7%;  Score 41;  DB 21;  Length 534;
  Best Local Similarity   75.0%;  Pred. No. 8e+02;
  Matches    6;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 HLMPFPGP 8
          ||||:| |
Db    260 HLMPYPAP 267


RESULT   7
US-60-167-217-7726
; Sequence 7726, Application US/60167217
; GENERAL INFORMATION:
;  APPLICANT: Li, Peter W. D.
;  TITLE OF INVENTION: ISOLATED DROSOPHILA PROTEINS, NUCLEIC
;  TITLE OF INVENTION: ACID MOLEUCLES ENCODING DROSOPHILA PROTEINS, AND USES
;  TITLE OF INVENTION: THEREOF
;  FILE REFERENCE: CL000152
;  CURRENT APPLICATION NUMBER: US/60/167,217
;  CURRENT FILING DATE: 1999-11-24
;  NUMBER OF SEQ ID NOS: 23195
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7726
;   LENGTH: 1021
;   TYPE: PRT
;   ORGANISM: Drosophila
US-60-167-217-7726


  Query Match             70.7%;  Score 41;  DB 23;  Length 1021;
  Best Local Similarity   75.0%;  Pred. No. 1.5e+03;
  Matches    6;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      3 MPFPGPLL 10
          :||||||:
Db    263 LPFPGPLM 270


RESULT   8
US-60-173-464-6181
; Sequence 6181, Application US/60173464
; GENERAL INFORMATION:
;  APPLICANT: Li, Peter W.D.
;  TITLE OF INVENTION: ISOLATED G-PROTEIN COUPLED RECEPTORS,
;  TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING GPCR PROTEINS AND USES
;  TITLE OF INVENTION: THEREOF
;  FILE REFERENCE: CL000173
;  CURRENT APPLICATION NUMBER: US/60/173,464
;  CURRENT FILING DATE: 1999-12-29
;  NUMBER OF SEQ ID NOS: 30269
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6181
;   LENGTH: 1060
;   TYPE: PRT
;   ORGANISM: Drosophila
US-60-173-464-6181


  Query Match             70.7%;  Score 41;  DB 23;  Length 1060;
  Best Local Similarity   75.0%;  Pred. No. 1.6e+03;
  Matches    6;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      3 MPFPGPLL 10
          :||||||:
Db    302 LPFPGPLM 309


RESULT   9
US-60-242-679-964
; Sequence 964, Application US/60242679
; GENERAL INFORMATION:
;  APPLICANT:  Ladunga, Steven Istvan
;  APPLICANT:  Spier, Eugene
;  APPLICANT:  Greenberg, Simon
;  APPLICANT:  Brandenberger, Ralph
;  APPLICANT:  Wang, Yu
;  APPLICANT:  Dubman, Alex
;  TITLE OF INVENTION:  ISOLATED HUMAN SECRETED PROTEINS,
;  TITLE OF INVENTION:  NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS, AND
;  TITLE OF INVENTION:  USES THEREOF
;  FILE REFERENCE:  CL000898-PROV
;  CURRENT APPLICATION NUMBER: US/60/242,679
;  CURRENT FILING DATE:  2000-10-24
;  NUMBER OF SEQ ID NOS:  2265
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 964
;   LENGTH: 265
;   TYPE: PRT
;   ORGANISM: HUMAN
US-60-242-679-964


  Query Match             69.0%;  Score 40;  DB 23;  Length 265;
  Best Local Similarity   75.0%;  Pred. No. 5.7e+02;
  Matches    6;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 HLMPFPGP 8
          ||:| |||
Db    106 HLLPVPGP 113


RESULT  10
US-60-167-217-7501
; Sequence 7501, Application US/60167217
; GENERAL INFORMATION:
;  APPLICANT: Li, Peter W. D.
;  TITLE OF INVENTION: ISOLATED DROSOPHILA PROTEINS, NUCLEIC
;  TITLE OF INVENTION: ACID MOLEUCLES ENCODING DROSOPHILA PROTEINS, AND USES
;  TITLE OF INVENTION: THEREOF
;  FILE REFERENCE: CL000152
;  CURRENT APPLICATION NUMBER: US/60/167,217
;  CURRENT FILING DATE: 1999-11-24
;  NUMBER OF SEQ ID NOS: 23195
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7501
;   LENGTH: 345
;   TYPE: PRT
;   ORGANISM: Drosophila
US-60-167-217-7501
```

```
RESULT   2
US-09-519-012-3
; Sequence 3, Application US/09519012
; GENERAL INFORMATION:
;  APPLICANT: Tang, Y. Tom
;  APPLICANT: Corley, Neil C.
;  APPLICANT: Lal, Preeti
;  APPLICANT: Kaser, Matthew R.
;  APPLICANT: Pearson, Cecelia
;  TITLE OF INVENTION: HUMAN GATA ZINC FINGER PROTEIN AND VARIANTS
;  FILE REFERENCE: PF-0474-1 CIP
;  CURRENT APPLICATION NUMBER: US/09/519,012
;  CURRENT FILING DATE: 2000-03-03
;  EARLIER APPLICATION NUMBER: 09/024,808
;  EARLIER FILING DATE: February 17, 1998
;  NUMBER OF SEQ ID NOS: 22
;  SOFTWARE: PERL Program
; SEQ ID NO 3
;   LENGTH: 413
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE: -
;   OTHER INFORMATION: g120956
US-09-519-012-3

  Query Match             100.0%;  Score 58;  DB 19;  Length 413;
  Best Local Similarity   100.0%;  Pred. No. 1.5;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 HLMPFPGPLL 10
          ||||||||||
Db    378 HLMPFPGPLL 387

RESULT   3
PCT-US01-01354-15768
; Sequence 15768, Application PC/TUS0101354
; GENERAL INFORMATION:
;  APPLICANT: Human Genome Sciences, Inc. et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PC004PCT
;  CURRENT APPLICATION NUMBER: PCT/US01/01354
;  CURRENT FILING DATE:  2001-03-17
;  NUMBER OF SEQ ID NOS: 42506
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15768
;   LENGTH: 98
;   TYPE: PRT
;   ORGANISM: Homo sapiens
PCT-US01-01354-15768

  Query Match             75.9%;  Score 44;  DB 1;  Length 98;
  Best Local Similarity   60.0%;  Pred. No. 52;
  Matches    6;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 HLMPFPGPLL 10
          |::||| |:|
Db     42 HVLPFPSPML 51

RESULT   4
US-09-252-991A-27500
; Sequence 27500, Application US/09252991A
; GENERAL INFORMATION:
;  APPLICANT:  Marc J. Rubenfield et al.
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;  TITLE OF INVENTION:  AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  107196.136
;  CURRENT APPLICATION NUMBER: US/09/252,991A
;  CURRENT FILING DATE:  1999-02-18
;  PRIOR APPLICATION NUMBER:  US 60/074,788
;  PRIOR FILING DATE:  1998-02-18
;  PRIOR APPLICATION NUMBER:  US 60/094,190
;  PRIOR FILING DATE:  1998-07-27
;  NUMBER OF SEQ ID NOS:  33142
; SEQ ID NO 27500
;   LENGTH: 406
;   TYPE: PRT
;   ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-27500

  Query Match             72.4%;  Score 42;  DB 16;  Length 406;
  Best Local Similarity   77.8%;  Pred. No. 4.3e+02;
  Matches    7;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 HLMPFPGPL 9
          | || ||||
Db    305 HKMPIPGPL 313

RESULT   5
PCT-US99-05804-111
; Sequence 111, Application PC/TUS9905804
; GENERAL INFORMATION:
;  APPLICANT: Human Genome Sciences, Inc., et al.
;  TITLE OF INVENTION: 95 Human Secreted Proteins
;  FILE REFERENCE: PZ027PCT
;  CURRENT APPLICATION NUMBER: PCT/US99/05804
;  CURRENT FILING DATE: 1999-03-18
;  EARLIER APPLICATION NUMBER: 60/078,566
;  EARLIER FILING DATE: 1998-03-19
;  EARLIER APPLICATION NUMBER: 60/078,574
;  EARLIER FILING DATE: 1998-03-19
;  EARLIER APPLICATION NUMBER: 60/078,576
;  EARLIER FILING DATE: 1998-03-19
;  EARLIER APPLICATION NUMBER: 60/078,563
;  EARLIER FILING DATE: 1998-03-19
;  EARLIER APPLICATION NUMBER: 60/078,573
;  EARLIER FILING DATE: 1998-03-19
;  EARLIER APPLICATION NUMBER: 60/078,578
;  EARLIER FILING DATE: 1998-03-19
;  EARLIER APPLICATION NUMBER: 60/078,579
;  EARLIER FILING DATE: 1998-03-19
;  EARLIER APPLICATION NUMBER: 60/078,581
;  EARLIER FILING DATE: 1998-03-19
;  EARLIER APPLICATION NUMBER: 60/078,577
;  EARLIER FILING DATE: 1998-03-19
;  EARLIER APPLICATION NUMBER: 60/080,314
;  EARLIER FILING DATE: 1998-04-01
;  EARLIER APPLICATION NUMBER: 60/080,312
;  EARLIER FILING DATE: 1998-04-01
;  EARLIER APPLICATION NUMBER: 60/080,313
;  EARLIER FILING DATE: 1998-04-01
;  NUMBER OF SEQ ID NOS: 392
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 111
;   LENGTH: 114
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (38)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (114)
;   OTHER INFORMATION: Xaa equals stop translation
PCT-US99-05804-111

  Query Match             70.7%;  Score 41;  DB 1;  Length 114;
  Best Local Similarity   75.0%;  Pred. No. 1.7e+02;
  Matches    6;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
                         GenCore version 4.5
              Copyright (c) 1993 - 2000  Compugen Ltd.

OM protein - protein search, using sw model

Run on:          October 23, 2001, 13:27:00 ; Search time 241.8 Seconds
                                              (without alignments)
                                              9.175 Million cell updates/sec

Title:           US-09-625-963-3
Perfect score:   58
Sequence:        1 HLMPFPGPLL 10

Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched:        2803329 seqs, 221847457 residues

Total number of hits satisfying chosen parameters:      2803329

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       Pending_Patents_AA_Main:*
               1:  /cgn2_6/ptodata/2/paa/PCTUS_COMB.pep:*
               2:  /cgn2_6/ptodata/2/paa/US06_COMB.pep:*
               3:  /cgn2_6/ptodata/2/paa/US07_COMB.pep:*
               4:  /cgn2_6/ptodata/2/paa/US080_COMB.pep:*
               5:  /cgn2_6/ptodata/2/paa/US081_COMB.pep:*
               6:  /cgn2_6/ptodata/2/paa/US082_COMB.pep:*
               7:  /cgn2_6/ptodata/2/paa/US083_COMB.pep:*
               8:  /cgn2_6/ptodata/2/paa/US084_COMB.pep:*
               9:  /cgn2_6/ptodata/2/paa/US085_COMB.pep:*
              10:  /cgn2_6/ptodata/2/paa/US086_COMB.pep:*
              11:  /cgn2_6/ptodata/2/paa/US087_COMB.pep:*
              12:  /cgn2_6/ptodata/2/paa/US088_COMB.pep:*
              13:  /cgn2_6/ptodata/2/paa/US089_COMB.pep:*
              14:  /cgn2_6/ptodata/2/paa/US090_COMB.pep:*
              15:  /cgn2_6/ptodata/2/paa/US091_COMB.pep:*
              16:  /cgn2_6/ptodata/2/paa/US092_COMB.pep:*
              17:  /cgn2_6/ptodata/2/paa/US093_COMB.pep:*
              18:  /cgn2_6/ptodata/2/paa/US094_COMB.pep:*
              19:  /cgn2_6/ptodata/2/paa/US095_COMB.pep:*
              20:  /cgn2_6/ptodata/2/paa/US096_COMB.pep:*
              21:  /cgn2_6/ptodata/2/paa/US097_COMB.pep:*
              22:  /cgn2_6/ptodata/2/paa/US098_COMB.pep:*
              23:  /cgn2_6/ptodata/2/paa/US60_COMB.pep:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                                  SUMMARIES
                      %
Result              Query
  No.    Score   Match Length DB  ID                       Description
------------------------------------------------------------------------
    1      58   100.0     10  20  US-09-625-963-3           Sequence 3, Appli
    2      58   100.0    413  19  US-09-519-012-3           Sequence 3, Appli
    3      44    75.9     98   1  PCT-US01-01354-15768     Sequence 15768, A
    4      42    72.4    406  16  US-09-252-991A-27500      Sequence 27500, A
    5      41    70.7    114   1  PCT-US99-05804-111       Sequence 111, App
    6      41    70.7    534  21  US-09-739-449-9641        Sequence 9641, Ap
    7      41    70.7   1021  23  US-60-167-217-7726        Sequence 7726, Ap
    8      41    70.7   1060  23  US-60-173-464-6181        Sequence 6181, Ap
    9      40    69.0    265  23  US-60-242-679-964         Sequence 964, App
   10      40    69.0    345  23  US-60-167-217-7501        Sequence 7501, Ap
   11      40    69.0    345  23  US-60-191-637-7407        Sequence 7407, Ap
   12      40    69.0    449  15  US-09-134-000-4944        Sequence 4944, Ap
   13      39    67.2     19  19  US-09-527-522-5           Sequence 5, Appli
   14      39    67.2     77  23  US-60-185-361-923         Sequence 923, App
   15      39    67.2    103  18  US-09-417-507-40482       Sequence 40482, A
   16      39    67.2    310  16  US-09-252-991A-27948      Sequence 27948, A
   17      38    65.5     29  14  US-09-008-186C-48         Sequence 48, Appl
   18      38    65.5     29  17  US-09-359-300-14          Sequence 14, Appl
   19      38    65.5     50  16  US-09-215-435-154         Sequence 154, App
   20      38    65.5     50  19  US-09-580-341-378         Sequence 378, App
   21      38    65.5     50  23  US-60-197-873-19427       Sequence 19427, A
   22      38    65.5    112   1  PCT-US98-12125-130       Sequence 130, App
   23      38    65.5    112  16  US-09-209-462B-135        Sequence 135, App
   24      38    65.5    114   1  PCT-US01-01354-12711     Sequence 12711, A
   25      38    65.5    117   1  PCT-US00-26323-70        Sequence 70, Appl
   26      38    65.5    200  18  US-09-417-507-31611       Sequence 31611, A
   27      38    65.5    340  19  US-09-515-038-60          Sequence 60, Appl
   28      38    65.5    360  17  US-09-372-348-9           Sequence 9, Appli
   29      38    65.5    399  23  US-60-253-625-2743        Sequence 2743, Ap
   30      38    65.5    399  23  US-60-257-931-3571        Sequence 3571, Ap
   31      38    65.5    399  23  US-60-269-308-4593        Sequence 4593, Ap
   32      38    65.5    452  16  US-09-252-991A-24195      Sequence 24195, A
   33      38    65.5    550  19  US-09-515-038-1           Sequence 1, Appli
   34      38    65.5    609  16  US-09-252-991A-24893      Sequence 24893, A
   35    37.5    64.7     45  17  US-09-307-140-1089        Sequence 1089, Ap
   36    37.5    64.7     45  22  US-09-817-076-1089        Sequence 1089, Ap
   37      37    63.8     12  13  US-08-926-808A-5          Sequence 5, Appli
   38      37    63.8     12  13  US-08-926-808C-5          Sequence 5, Appli
   39      37    63.8    110  23  US-60-161-932-1567        Sequence 1567, Ap
   40      37    63.8    213  18  US-09-417-507-31491       Sequence 31491, A
   41      37    63.8    239  18  US-09-417-507-41630       Sequence 41630, A
   42      37    63.8    313  23  US-60-205-306-24          Sequence 24, Appl
   43      37    63.8    443  16  US-09-252-991A-30036      Sequence 30036, A
   44      37    63.8    778  12  US-08-808-599-5           Sequence 5, Appli
   45      37    63.8   1069   1  PCT-US00-10644-10        Sequence 10, Appl

                                 ALIGNMENTS


RESULT   1
US-09-625-963-3
; Sequence 3, Application US/09625963
; GENERAL INFORMATION:
;  APPLICANT:  Strauss, Hans Josef
;  APPLICANT:  Gao, Liquan
;  TITLE OF INVENTION:  Immunotherapeutic Methods Using Epitopes of WT-1 and
;  TITLE OF INVENTION:  GATA-1
;  FILE REFERENCE:  ICI 101
;  CURRENT APPLICATION NUMBER: US/09/625,963
;  CURRENT FILING DATE:  2000-07-26
;  PRIOR APPLICATION NUMBER:  PCT/GB99/03572
;  PRIOR FILING DATE:  1999-11-02
;  PRIOR APPLICATION NUMBER:  GB9823897.5
;  PRIOR FILING DATE:  1998-11-02
;  NUMBER OF SEQ ID NOS:  5
;  SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 3
;   LENGTH: 10
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-625-963-3


  Query Match             100.0%;  Score 58;  DB 20;  Length 10;
  Best Local Similarity   100.0%;  Pred. No. 0.038;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 HLMPFPGPLL 10
              ||||||||||
Db          1 HLMPFPGPLL 10
```

```
  Query Match             63.8%;  Score 37;  DB 17;  Length 778;
  Best Local Similarity   66.7%;  Pred. No. 2.2e+02;
  Matches    6;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 HLMPFPGPL 9
          |||| | |:
Db    303 hlmpspapv 311


RESULT  14
AAB23359
ID   AAB23359 standard; Protein; 778 AA.
XX
AC   AAB23359;
XX
DT   09-JAN-2001  (first entry)
XX
DE   Human tastin protein.
XX
KW   Trophinin; vaccine; contraceptive; mouse; human; tastin.
XX
OS   Homo sapiens.
XX
PN   US6111089-A.
XX
PD   29-AUG-2000.
XX
PF   28-FEB-1997;   97US-0808599.
XX
PR   04-OCT-1994;   94US-0317522.
PR   12-MAY-1995;   95US-0439818.
XX
PA   (BURN-) BURNHAM INST.
XX
PI   Fukuda MN;
XX
DR   WPI; 2000-586380/55.
DR   N-PSDB; AAA91965.
XX
PT   Novel nucleic acid sequence encoding mouse trophinin, for producing
PT   trophinin vaccine useful to prevent pregnancy in an individual by
PT   eliciting an immune response which blocks embryo implantation  -
XX
PS   Disclosure; Column 57-62; 57pp; English.
XX
CC   The present invention relates to mouse trophinin sequence (see
CC   AAB23367). Mouse trophinin is used as a vaccine for preventing
CC   pregnancy in an individual by eliciting an immune response against
CC   trophinin that blocks embryo implantation. Due to the limited
CC   expression of trophinin, an anti-trophinin immune response is less
CC   toxic than other prior art anti-pregnancy vaccines directed to peptide
CC   hormones. The present sequence is human tastin. Tastin is a
CC   trophinin-assisting protein.
XX
SQ   Sequence   778 AA;

  Query Match             63.8%;  Score 37;  DB 21;  Length 778;
  Best Local Similarity   66.7%;  Pred. No. 2.2e+02;
  Matches    6;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 HLMPFPGPL 9
          |||| | |:
Db    303 hlmpspapv 311


RESULT  15
AAB36839
ID   AAB36839 standard; Protein; 1069 AA.
XX
AC   AAB36839;
XX
DT   20-FEB-2001  (first entry)
XX
DE   Protein sequence associated with IRRR.
XX
KW   Insulin receptor-related receptor; IRRR; chromosome 1q21-q24;
KW   obesity; dyslipidemia; diabetes.
XX
OS   Rattus rattus.
XX
PN   WO200065090-A2.
XX
PD   02-NOV-2000.
XX
PF   19-APR-2000; 2000WO-US10644.
XX
PR   22-APR-1999;   99US-0296906.
PR   22-JUN-1999;   99US-0337976.
XX
PA   (ZYMO ) ZYMOGENETICS INC.
XX
PI   Lok S,  Whitmore TE;
XX
DR   WPI; 2000-687365/67.
XX
PT   Detecting a chromosome 1q21-q24 abnormality for diagnosing metabolic
PT   disease, such as human obesity and diabetic disorders, comprises
PT   examining insulin receptor-related receptor gene and its gene products
PT   -
XX
PS   Claim 21; Page 94-97; 111pp; English.
XX
CC   The present invention relates to insulin receptor-related receptor
CC   (IRRR). Mutations in this gene indicate a chromosome 1q21-q24
CC   abnormality. IRRR polypeptides and DNA may be useful in the diagnosis of
CC   of disorders associated with abnormal expression of the IRRR protein,
CC   for example obesity, dyslipidemia and diabetes.
XX
SQ   Sequence   1069 AA;

  Query Match             63.8%;  Score 37;  DB 21;  Length 1069;
  Best Local Similarity   60.0%;  Pred. No. 3.1e+02;
  Matches    6;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 HLMPFPGPLL 10
          |  :|:| |||
Db    406 hhlpwapll 415


Search completed: October 23, 2001, 13:21:03
Job time: 56 sec
```

```
XX
PT   New protease - which cleaves at cysteines
XX
PS   Claim 1; Page 14; 32pp; Japanese.
XX
CC   This invention describes novel proteins and their encoding nucleic
CC   acids isolated from Sarcophaga peregrina which have cysteine
CC   protease-like activity.
XX
SQ   Sequence   550 AA;


  Query Match             65.5%;  Score 38;  DB 20;  Length 550;
  Best Local Similarity   100.0%;  Pred. No. 1.1e+02;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      3 MPFPGP 8
          ||||||
Db    214 mpfpgp 219


RESULT 13
AAR94900
ID   AAR94900 standard; Protein; 778 AA.
XX
AC   AAR94900;
XX
DT   03-JUL-1996  (first entry)
XX
DE   Human tastin.
XX
KW   Trophinin; trophinin-assisting protein; tastin; bystin; lastin;
KW   embryo implantation; infertility; cell adhesion; therapy; diagnosis;
KW   tumour suppressor; cancer.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Region          516..650
FT                   /note= "cystine-rich 4-tandem repeat region"
FT   Modified-site   2
FT                   /note= "protein kinase C phosphorylation site"
FT   Modified-site   26
FT                   /note= "protein kinase C phosphorylation site"
FT   Modified-site   98
FT                   /note= "protein kinase C phosphorylation site"
FT   Modified-site   129
FT                   /note= "casein kinase II phosphorylation site"
FT   Modified-site   135
FT                   /note= "protein kinase C phosphorylation site"
FT   Modified-site   155
FT                   /note= "protein kinase C phosphorylation site"
FT   Modified-site   161
FT                   /note= "protein kinase C phosphorylation site"
FT   Modified-site   166
FT                   /note= "protein kinase C phosphorylation site"
FT   Modified-site   177
FT                   /note= "MAP kinase phosphorylation site"
FT   Modified-site   179
FT                   /note= "protein kinase C phosphorylation site"
FT   Modified-site   191
FT                   /note= "casein kinase II phosphorylation site"
FT   Modified-site   234
FT                   /note= "cAMP/cGMP-dependent phosphorylation site"
FT   Modified-site   236
FT                   /note= "protein kinase C phosphorylation site"
FT   Modified-site   237
FT                   /note= "protein kinase C phosphorylation site"
FT   Modified-site   269
FT                   /note= "casein kinase II phosphorylation site"
FT   Modified-site   274
FT                   /note= "protein kinase C phosphorylation site"
FT   Modified-site   296
FT                   /note= "casein kinase II phosphorylation site"
FT   Modified-site   350
FT                   /label= cAMP/cGMP-dependent phosphorylation site
FT   Modified-site   363
FT                   /note= "MAP kinase phosphorylation site"
FT   Modified-site   379
FT                   /note= "casein kinase II phosphorylation site"
FT   Modified-site   408
FT                   /note= "casein kinase II phosphorylation site"
FT   Modified-site   421
FT                   /note= "protein kinase C phosphorylation site"
FT   Modified-site   424
FT                   /note= "casein kinase II phosphorylation site"
FT   Modified-site   426
FT                   /note= "casein kinase II phosphorylation site"
FT   Modified-site   452
FT                   /note= "protein kinase C phosphorylation site"
FT   Modified-site   478
FT                   /note= "casein kinase II phosphorylation site"
FT   Modified-site   569
FT                   /note= "protein kinase C phosphorylation site"
FT   Modified-site   570
FT                   /note= "casein kinase II phosphorylation site"
FT   Modified-site   602
FT                   /note= "protein kinase C phosphorylation site"
FT   Modified-site   603
FT                   /note= "casein kinase II phosphorylation site"
FT   Modified-site   691
FT                   /note= "protein kinase C phosphorylation site"
FT   Modified-site   714
FT                   /note= "protein kinase C phosphorylation site"
FT   Modified-site   737
FT                   /note= "casein kinase II phosphorylation site"
FT   Modified-site   738
FT                   /note= "protein kinase C phosphorylation site"
FT   Modified-site   756
FT                   /note= "casein kinase II phosphorylation site"
XX
PN   WO9610414-A1.
XX
PD   11-APR-1996.
XX
PF   04-OCT-1995;   95WO-US13259.
XX
PR   12-MAY-1995;   95US-0439818.
PR   04-OCT-1994;   94US-0317522.
XX
PA   (LJOL-) LA JOLLA CANCER RES FOUND.
XX
PI   Fukuda MN;
XX
DR   WPI; 1996-209192/21.
DR   N-PSDB; AAT18677.
XX
PT   Mammalian trophinin and trophinin-assisting protein - used in
PT   inhibiting or enhancing embryo implantation, diagnosis of
PT   infertility and treatment of cancer
XX
PS   Claim 24; Fig 6; 106pp; English.
XX
CC   Human trophinin-assisting proteins tastin (AAR94900), bystin (AAR94902)
CC   and lastin (AAR94903) are the products of 3 cDNA clones (AAT18677-79,
CC   respectively).  Expression of these proteins is required by a cell
CC   to effect trophinin-mediated cell adhesion.  The trophinin-assisting
CC   proteins can function to segregate trophinin molecules (see also
CC   AAR94895) into clusters on the apical plasma membrane.  They can be
CC   used to improve trophinin-mediated cell adhesion, e.g. to minimise
CC   embryo implantation failure.  Tastin is probably a cytoplasmic
CC   protein and shows no overall significant homology to reported
CC   protein sequences.
XX
SQ   Sequence   778 AA;
```

```
PF   06-JUL-1999;   99WO-US12366.
XX
PR   06-JUL-1998;   98US-0110938.
PR   13-JUL-1998;   98US-0114466.
PR   23-JUL-1998;   98US-0093897.
PR   12-AUG-1998;   98US-0132968.
PR   18-AUG-1998;   98US-0136214.
PR   11-SEP-1998;   98US-0099999.
XX
PA   (SCHE ) SCHERING CORP.
XX
PI   Bates EEM,  Lebecque SJE,  Murphy EE,  Mattson JD,  Gorman DM;
PI   Hedrick JA,  Wang L,  Zlotnik A,  Murgolo NJ,  Greene JR,  Johnston JA;
PI   Bazan JF,  Mahony D,  Lees EM;
XX
DR   WPI; 2000-171015/15.
DR   N-PSDB; AAZ92413.
XX
PT   New isolated mammalian genes, used to develop products for treating
PT   e.g. immune, inflammatory or allergic abnormalities, cancers or
PT   degenerative conditions  -
XX
PS   Claim 43 ; Page 184-185; 218pp; English.
XX
CC   The invention relates to a number of primate and/or rodent proteins,.and
CC   the genes which encode them. The invention encompasses human dendritic
CC   cell prostaglandin transporter (DC-PGT); the TNF (tumour necrosis
CC   factor) receptor family-related proteins HDTEA84, HSLJD37R and RANKL;
CC   human CC chemokine HCC5; human deubiquitinating proteins Dub11 and Dub
CC   12; human MD-1 and human and murine MD-2 proteins, which exhibit the
CC   properties of ligands for proteins comprising a leucine-rich motif
CC   (LRR); human cyclin E2; cDNAs encoding these proteins; and antibodies
CC   against these proteins. The proteins can be used for modulating the
CC   physiology or development of a cell. They can be used to mediate uptake
CC   of substrates (e.g., prostaglandin-like molecules), to modulate or
CC   mediate cellular interactions (e.g., induce or prevent trafficking,
CC   proliferation, or differentiation of cells), or are intracellular
CC   proteins which are important in various cellular processes such as the
CC   deubiquitination of proteins or cell cycle regulation. The products can
CC   be used for treating medical conditions such as immune, inflammatory or
CC   allergic disorders, or abnormal cellular proliferation, for example,
CC   cancers or degenerative conditions. They can be used to modulate immune
CC   responses in disease states e.g., autoimmune disorders, including
CC   rheumatoid arthritis, systemic lupus erythematosus, Hashimoto's
CC   autoimmune thyroiditis, as well as acute and chronic inflammatory
CC   responses in which T cell activation, expansion, and/or immunological T
CC   cell memory play an important role. Sequences AAY77470-Y77473 represent
CC   human deubiquitinating proteins (Dub). AAY77470-Y77471 encode Dub11
CC   proteins and AAY77472-Y77473 encode Dub12 proteins.
XX
SQ   Sequence   360 AA;

  Query Match             65.5%;  Score 38;  DB 21;  Length 360;
  Best Local Similarity   100.0%;  Pred. No. 71;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 MPFPGP 8
              ||||||
Db          1 mpfpgp 6


RESULT  11
AAR82242
ID   AAR82242 standard; Protein; 498 AA.
XX
AC   AAR82242;
XX
DT   10-APR-1996  (first entry)
XX
DE   Capsanthin-capsorubin synthase from Capsicum annuum.
XX
KW   xanthophyll biosynthesis; Capsicum; pepper; red food colourant;
KW   5,6-epoxycarotenoid; bifunctional enzyme; CCS; chromoplast membrane.
XX
OS   Capsicum annuum.
XX
PN   WO9523863-A1.
XX
PD   08-SEP-1995.
XX
PF   17-FEB-1995;   95WO-EP00584.
XX
PR   23-MAR-1994;   94EP-0400626.
PR   01-MAR-1994;   94GB-0003943.
XX
PA   (CNRS ) CENT NAT RECH SCI.
XX
PI   Camara B,  Kuntz M;
XX
DR   WPI; 1995-320578/41.
DR   N-PSDB; AAT03851.
XX
PT   DNA construct for modifying carotenoid production in plants -
PT   comprises sequences homologous to xanthophyll biosynthetic or
PT   degradative enzyme encoding sequences
XX
PS   Claim 6; Fig 3; 49pp; English.
XX
CC   The capsanthin-capsorubin synthase (CCS) enzyme was isolated from
CC   the chromoplast membrane fraction of Capsicum annuum (pepper). The
CC   bifunctional enzyme catalyses conversion of the ubiquitous 5,6-epoxy
CC   carotenoids, antheraxanthin and violaxanthin, into capsanthin and
CC   capsorubin, respectively. These xanthophylls are used as food
CC   colourants. Also the colour of petals, leaves, fruits, etc. of
CC   plants transformed by DNA coding for the CCS enzyme can be modified
CC   by prodn. of the xanthophylls.
XX
SQ   Sequence   498 AA;

  Query Match             65.5%;  Score 38;  DB 16;  Length 498;
  Best Local Similarity   77.8%;  Pred. No. 98;
  Matches    7;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 LMPFPGPLL 10
              | ||| |||
Db          5 lkpfpspll 13


RESULT  12
AAY08899
ID   AAY08899 standard; protein; 550 AA.
XX
AC   AAY08899;
XX
DT   20-AUG-1999  (first entry)
XX
DE   S. peregrina cysteine protease-like protein 1.
XX
KW   Cysteine protease; cysteine protease-like.
XX
OS   Sarcophaga peregrina.
XX
PN   JP11146789-A.
XX
PD   02-JUN-1999.
XX
PF   18-NOV-1997;   97JP-0333474.
XX
PR   18-NOV-1997;   97JP-0333474.
XX
PA   (UYTY ) UNIV TOKYO.
XX
DR   WPI; 1999-378998/32.
DR   N-PSDB; AAX78033.
```

```
CC   proteins obtained may have cytokine activity, cell
CC   proliferation/differentiation activity, haematopoiesis regulating
CC   activity, tissue growth regulating activity, reproductive hormone
CC   regulating activity, chemotactic/ chemokinetic activity, haemostatic and
CC   thrombolytic activity, receptor/ ligand activity, anti-inflammatory
CC   activity, tumour inhibition activity or other activities. The products
CC   can be used in forensic, gene therapy and chromosome mapping procedures.
CC   The sequences can also be used for obtaining corresponding promoter
CC   sequences. The nucleic acids encoding the signal peptide can be used for
CC   directing extracellular secretion of a polypeptide or the insertion of a
CC   polypeptide into a membrane, or importing a polypeptide into a cell.
XX
SQ   Sequence   50 AA;


  Query Match             65.5%;  Score 38;  DB 20;  Length 50;
  Best Local Similarity   70.0%;  Pred. No. 10;
  Matches    7;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 HLMPFPGPLL 10
                ||  |||| |
Db         20 hlasfpgpsl 29


RESULT   9
AAW78145
ID   AAW78145 standard; Protein; 112 AA.
XX
AC   AAW78145;
XX
DT   13-APR-1999  (first entry)
XX
DE   Human secreted protein encoded by gene 20 clone HSKZE52.
XX
KW   Human; secreted protein; fusion protein; gene therapy; protein therapy;
KW   diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;
KW   developmental abnormality; foetal deficiency; blood; allergy; renal;
KW   immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
KW   inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;
KW   cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
KW   osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
KW   endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 112
FT                   /label= unknown
XX
PN   WO9856804-A1.
XX
PD   17-DEC-1998.
XX
PF   11-JUN-1998;   98WO-US12125.
XX
PR   02-OCT-1997;   97US-0061060.
PR   13-JUN-1997;   97US-0049547.
PR   13-JUN-1997;   97US-0049548.
PR   13-JUN-1997;   97US-0049549.
PR   13-JUN-1997;   97US-0049550.
PR   13-JUN-1997;   97US-0049606.
PR   13-JUN-1997;   97US-0049607.
PR   13-JUN-1997;   97US-0049608.
PR   13-JUN-1997;   97US-0049609.
PR   13-JUN-1997;   97US-0049610.
PR   13-JUN-1997;   97US-0049611.
PR   13-JUN-1997;   97US-0050566.
PR   13-JUN-1997;   97US-0050901.
PR   13-JUN-1997;   97US-0052989.
PR   08-JUL-1997;   97US-0051919.
PR   18-AUG-1997;   97US-0055984.
PR   12-SEP-1997;   97US-0058665.
PR   12-SEP-1997;   97US-0058668.
PR   12-SEP-1997;   97US-0058669.
PR   12-SEP-1997;   97US-0058750.
PR   12-SEP-1997;   97US-0058971.
PR   12-SEP-1997;   97US-0058972.
PR   12-SEP-1997;   97US-0058975.
PR   02-OCT-1997;   97US-0060834.
PR   02-OCT-1997;   97US-0060841.
PR   02-OCT-1997;   97US-0060844.
PR   02-OCT-1997;   97US-0060865.
PR   02-OCT-1997;   97US-0061059.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Brewer LA,  Ebner R,  Ferrie AM,  Feng P,  Greene JM,  Lafleur DW;
PI   Moore PA,  Ni J,  Olsen HS,  Rosen CA,  Ruben SM,  Shi Y,  Young P;
PI   Yu GL;
XX
DR   WPI; 1999-080881/07.
DR   N-PSDB; AAX04330.
XX
PT   New isolated human genes and the secreted polypeptides they encode -
PT   useful for diagnosis and treatment of e.g. cancers, neurological
PT   disorders, immune diseases, inflammation or blood disorders
XX
PS   Claim 11; Page 272; 380pp; English.
XX
CC   This sequence represents a secreted human protein encoded by the gene
CC   clone detailed in the descriptor line.  The gene can be used to generate
CC   fusion proteins by linking to the gene to a human immunoglobulin Fc
CC   portion (e.g. AAX04302) for increasing the stability of the fused
CC   protein as compared to the human protein only.
CC   The invention relates to 86 novel genes and their fragments (nucleic
CC   acid sequences: AAX04311-X04410; amino acid sequences AAW78126-W78225)
CC   which are useful for preventing, treating or ameliorating  medical
CC   conditions e.g. by protein or gene therapy.  Also, pathological
CC   conditions can be diagnosed by determining the amount of the new
CC   polypeptides in a sample or by determining the presence of mutations in
CC   the new polynucleotides. Specific uses are described for each of the 86
CC   polynucleotides, based on which tissues they are most highly expressed in
CC   (see AAX04311 for described uses).
XX
SQ   Sequence   112 AA;


  Query Match             65.5%;  Score 38;  DB 20;  Length 112;
  Best Local Similarity   85.7%;  Pred. No. 22;
  Matches    6;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 LMPFPGP 8
                |:|||||
Db         77 lvpfpgp 83


RESULT  10
AAY77470
ID   AAY77470 standard; Protein; 360 AA.
XX
AC   AAY77470;
XX
DT   05-JUN-2000  (first entry)
XX
DE   Human deubiquitinating protein Dub11, SEQ ID NO:32.
XX
KW   Human; deubiquitinating protein; Dub11; Dub12; immune disorder;
KW   inflammation; allergy; immunosuppressant; antiarthritic; antirheumatoid;
KW   antiinflammatory; dermatological; antithyroid.
XX
OS   Homo sapiens.
XX
PN   WO200001817-A2.
XX
PD   13-JAN-2000.
XX
```

```
XX
PR   19-MAR-1999;   99JP-0075511.
PR   30-MAR-1999;   99JP-0088634.
XX
PA   (SUMO ) SUMITOMO CHEM CO LTD.
XX
PI   Takashima Y,  Mitsuda S,  Wieser M;
XX
DR   WPI; 2000-629576/61.
XX
PT   Novel Mycobacterium protein capable of catalyzing transamination stereo
PT   selectively, useful for producing optically active amino compounds,
PT   comprises a sequence of 339 amino acids   -
XX
PS   Example 10; Page 40; 54pp; English.
XX
CC   The present sequence is a peptide used to design PCR primers used during
CC   the identification of the coding sequence for a protein from
CC   Mycobacterium aurum which acts as a biocatalyst. The protein is able to
CC   catalyse transamination stereoselectively to produce an optically active
CC   amino compound.
XX
SQ   Sequence   19 AA;


  Query Match            67.2%;  Score 39;  DB 21;  Length 19;
  Best Local Similarity   70.0%;  Pred. No. 2.7;
  Matches    7;  Conservative    1;  Mismatches    2;  Indels  · 0;  Gaps    0;

Qy        1 HLMPFPGPLL 10
            |||  ||||:
Db        5 hlmdepgpli 14


RESULT   7
AAY35905
ID   AAY35905 standard; Protein; 50 AA.
XX
AC   AAY35905;
XX
DT   13-SEP-1999  (first entry)
XX
DE   Extended human secreted protein sequence, SEQ ID NO. 154.
XX
KW   Secreted protein; human; cytokine; cellular proliferation; cell movement;
KW   cellular differentiation; immune system regulator; anti-inflammatory;
KW   haematopoiesis regulator; tissue growth regulator; tumour inhibitor;
KW   reproductive hormone regulator; chemotaxis; chemokinesis; gene therapy;
KW   genetic disease.
XX
OS   Homo sapiens.
XX
PN   WO9931236-A2.
XX
PD   24-JUN-1999.
XX
PF   17-DEC-1998;   98WO-IB02122.
XX
PR   10-AUG-1998;   98US-0096116.
PR   17-DEC-1997;   97US-0069957.
PR   09-FEB-1998;   98US-0074121.
PR   13-APR-1998;   98US-0081563.
XX
PA   (GEST ) GENSET.
XX
PI   Bougueleret L,  Duclert A,  Dumas Milne Edwards J;
XX
DR   WPI; 1999-385906/32.
DR   N-PSDB; AAX97589.
XX
PT   New isolated human secreted proteins
XX
PS   Claim 9; Page 186-187; 516pp; English.
```

```
XX
CC   This sequence is encoded by an extended human secreted protein coding
CC   sequence of the invention. The secreted proteins can be used in treating
CC   or controlling a variety of human conditions. The secreted proteins may
CC   act as cytokines or may affect cellular proliferation or differentiation
CC   or may act as immune system regulators, haematopoiesis regulators, tissue
CC   growth regulators, regulators of reproductive hormones or cell movement
CC   or have chemotactic/chemokinetic, receptor/ligand, anti-inflammatory or
CC   tumour inhibition activity. The DNAs can be used in forensic procedures
CC   to identify individuals or in diagnostic procedures to identify
CC   individuals having genetic diseases resulting from abnormal expression of
CC   the genes corresponding to the extended cDNAs. They are also useful for
CC   constructing a high resolution map of the human chromosomes. They can
CC   also be used for gene therapy to control or treat genetic diseases.
XX
SQ   Sequence   50 AA;


  Query Match            65.5%;  Score 38;  DB 20;  Length 50;
  Best Local Similarity   70.0%;  Pred. No. 10;
  Matches    7;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy        1 HLMPFPGPLL 10
            ||  |||| |
Db       20 hlasfpgpsl 29


RESULT   8
AAY13094
ID   AAY13094 standard; Protein; 50 AA.
XX
AC   AAY13094;
XX
DT   22-JUN-1999  (first entry)
XX
DE   Human secreted protein encoded by 5' EST SEQ ID NO: 108.
XX
KW   Human; secreted protein; EST; expressed sequence tag; diagnosis;
KW   forensic; gene therapy; chromosome mapping; signal peptide;
KW   upstream regulatory sequence; cytokine activity; cell proliferation;
KW   differentiation; haematopoiesis regulation; tissue growth regulation;
KW   reproductive hormone regulation; chemotactic; chemokinetic; haemostatic;
KW   thrombolytic; anti-inflammatory; tumour inhibition.
XX
OS   Homo sapiens.
XX
PN   WO9906552-A2.
XX
PD   11-FEB-1999.
XX
PF   31-JUL-1998;   98WO-IB01236.
XX
PR   01-AUG-1997;   97US-0905223.
XX
PA   (GEST ) GENSET.
XX
PI   Duclert A,  Dumas Milne Edwards J,  Lacroix B;
XX
DR   WPI; 1999-153782/13.
DR   N-PSDB; AAX51894.
XX
PT   New isolated brain-derived nucleic acids - used to develop products
PT   which may have cytokine, immune, regulatory, haematopoiesis
PT   regulating, anti-inflammatory or tumour inhibition activity
XX
PS   Claim 34; Page 501; 577pp; English.
XX
CC   AAX51787 to AAX52019 represent 5' expressed sequence tags (ESTs) for
CC   human secreted proteins, and encode the proteins given in AAY12987 to
CC   AAY13219, respectively. The proteins given represent the signal peptide
CC   and an N-terminal fragment of a secreted protein. The nucleic acid
CC   sequences can be used for producing secreted human gene products. They
CC   can also be used to develop products for diagnosis and therapy. The
```

```
  Query Match             69.0%;  Score 40;  DB 18;  Length 11;
  Best Local Similarity   75.0%;  Pred. No. 1.1;
  Matches    6;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2 LMPFPGPL 9
           |:|||||:
Db       2 lvpfpgpi 9


RESULT   4
AAR53617
ID   AAR53617 standard; peptide; 9 AA.
XX
AC   AAR53617;
XX
DT   25-JAN-1995  (first entry)
XX
DE   Opioid peptide #11.
XX
KW   Opioid; activity; hydrolysis; lactoprotein; neutral protease;
KW   alkali protease; gastrointestinal hormone stimulating agents;
KW   analgesics; hypnotics; electrolyte adsorption stimulating agents;
KW   diarrhoea treating agents; drug.
XX
OS   Mammalian.
XX
PN   JP06128287-A.
XX
PD   10-MAY-1994.
XX
PF   02-NOV-1992;   92JP-0315515.
XX
PR   07-NOV-1991;   91JP-0318569.
PR   25-DEC-1991;   91JP-0356633.
PR   01-SEP-1992;   92JP-0255403.
XX
PA   (NISS ) NISSHIN FLOUR MILLING CO.
XX
DR   WPI; 1994-188987/23.
XX
PT   New opioid peptide(s) and their salts - are used as analgesic,
PT   for stimulating e.g. electrolyte absorption and treating
PT   diarrhoea
XX
PS   Disclosure; Page 13; 15pp; Japanese.
XX
CC   The sequences given in AAR53607-26 are peptides which have opioid
CC   activity.  These peptides are produced by hydrolysing lactoprotein
CC   with neutral protease derived from alkali protease.  The peptides and
CC   their salts are useful as analgesics, hypnotics, gastrointestinal
CC   hormone stimulating agents, electrolyte adsorption stimulating agents,
CC   diarrhoea treating agents, etc.  The peptide given in AAR53608 shows
CC   lower opioid activity and is therefore expected to be used as processed
CC   drug.
XX
SQ   Sequence   9 AA;

  Query Match             67.2%;  Score 39;  DB 15;  Length 9;
  Best Local Similarity   85.7%;  Pred. No. 3.4e+05;
  Matches    6;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       4 PFPGPLL 10
           |||||:|
Db       2 pfpgpil 8


RESULT   5
AAR53622
ID   AAR53622 standard; peptide; 10 AA.
XX
```

```
AC   AAR53622;
XX
DT   25-JAN-1995  (first entry)
XX
DE   Opioid peptide #16.
XX
KW   Opioid; activity; hydrolysis; lactoprotein; neutral protease;
KW   alkali protease; gastrointestinal hormone stimulating agents;
KW   analgesics; hypnotics; electrolyte adsorption stimulating agents;
KW   diarrhoea treating agents; drug.
XX
OS   Mammalian.
XX
PN   JP06128287-A.
XX
PD   10-MAY-1994.
XX
PF   02-NOV-1992;   92JP-0315515.
XX
PR   07-NOV-1991;   91JP-0318569.
PR   25-DEC-1991;   91JP-0356633.
PR   01-SEP-1992;   92JP-0255403.
XX
PA   (NISS ) NISSHIN FLOUR MILLING CO.
XX
DR   WPI; 1994-188987/23.
XX
PT   New opioid peptide(s) and their salts - are used as analgesic,
PT   for stimulating e.g. electrolyte absorption and treating
PT   diarrhoea
XX
PS   Disclosure; Page 13; 15pp; Japanese.
XX
CC   The sequences given in AAR53607-26 are peptides which have opioid
CC   activity.  These peptides are produced by hydrolysing lactoprotein
CC   with neutral protease derived from alkali protease.  The peptides and
CC   their salts are useful as analgesics, hypnotics, gastrointestinal
CC   hormone stimulating agents, electrolyte adsorption stimulating agents,
CC   diarrhoea treating agents, etc.  The peptide given in AAR53608 shows
CC   lower opioid activity and is therefore expected to be used as processed
CC   drug.
XX
SQ   Sequence   10 AA;

  Query Match             67.2%;  Score 39;  DB 15;  Length 10;
  Best Local Similarity   85.7%;  Pred. No. 1.4;
  Matches    6;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       4 PFPGPLL 10
           |||||:|
Db       3 pfpgpil 9


RESULT   6
AAB26103
ID   AAB26103 standard; Protein; 19 AA.
XX
AC   AAB26103;
XX
DT   15-JAN-2001  (first entry)
XX
DE   Stereoselective transamination catalyst gene PCR primer peptide #3.
XX
KW   Transamination; optically active amino compound; isomer; PCR primer.
XX
OS   Mycobacterium aurum.
XX
PN   EP1038953-A1.
XX
PD   27-SEP-2000.
XX
PF   14-MAR-2000; 2000EP-0400701.
```

```
CC   The present sequence is peptide epitope hug 378-87, produced by gata-1
CC   expressing cells and found at residues 378-387 of the gata-1 protein,
CC   which is recognised by cytotoxic T lymphocytes. Gata-1 is aberrantly
CC   expressed in leukaemias. The peptide can be used as a vaccine to
CC   stimulate the elimination, by cytotoxic T lymphocytes, of cancer cells
CC   aberrantly expressing gata-1. In addition, the nucleic acid encoding the
CC   peptide may also be used in the same manner. Alternatively, the peptide
CC   may be used in vitro to produce activated cytotoxic T lymphocytes.
XX
SQ   Sequence   10 AA;


  Query Match             100.0%;  Score 58;  DB 21;  Length 10;
  Best Local Similarity   100.0%;  Pred. No. 0.00095;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 HLMPFPGPLL 10
              ||||||||||
Db          1 hlmpfpgpll 10


RESULT   2
AAY41311
ID   AAY41311 standard; Protein; 114 AA.
XX
AC   AAY41311;
XX
DT   02-DEC-1999  (first entry)
XX
DE   Human secreted protein encoded by gene 4 clone HLDON23.
XX
KW   Human; secreted protein; fusion protein; gene therapy; protein therapy;
KW   diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;
KW   developmental abnormality; foetal deficiency; blood; allergy; renal;
KW   immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
KW   inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;
KW   cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
KW   osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
KW   endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.
XX
OS   Homo sapiens.
XX
PN   WO9947540-A1.
XX
PD   23-SEP-1999.
XX
PF   18-MAR-1999;   99WO-US05804.
XX
PR   19-MAR-1998;   98US-0078563.
PR   19-MAR-1998;   98US-0078566.
PR   19-MAR-1998;   98US-0078573.
PR   19-MAR-1998;   98US-0078574.
PR   19-MAR-1998;   98US-0078576.
PR   19-MAR-1998;   98US-0078577.
PR   19-MAR-1998;   98US-0078578.
PR   19-MAR-1998;   98US-0078579.
PR   19-MAR-1998;   98US-0078581.
PR   01-APR-1998;   98US-0080312.
PR   01-APR-1998;   98US-0080313.
PR   01-APR-1998;   98US-0080314.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Ruben SM,  Ni J,  Rosen CA,  Yu G,  Young PE,  Feng P,  Soppet DR;
PI   Wei Y,  Endress GA,  Duan RD,  Kyaw H,  Ebner R,  Lafleur DW;
PI   Olsen HS,  Shi Y,  Moore PA;
XX
DR   WPI; 1999-562050/47.
DR   N-PSDB; AAZ24814.
XX
PT   New isolated human genes, useful for diagnosis and treatment of e.g.
PT   cancers, neurological disorders, immune diseases, inflammation or blood
PT   disorders  -
XX
PS   Claim 11; Page 360; 484pp; English.
XX
CC   This sequence represents a secreted human protein encoded by the gene
CC   clone detailed in the descriptor line.  The gene can be used to generate
CC   fusion proteins by linking to the gene to a human immunoglobulin Fc
CC   portion (e.g. AAZ24802) for increasing the stability of the fused
CC   protein as compared to the human protein only.
CC   The invention relates to 95 novel genes and their fragments (nucleic
CC   acid sequences: AAZ24811-Z24907; amino acid sequences AAY41308-Y41404)
CC   which are useful for preventing, treating or ameliorating  medical
CC   conditions e.g. by protein or gene therapy.  Also, pathological
CC   conditions can be diagnosed by determining the amount of the new
CC   polypeptides in a sample or by determining the presence of mutations in
CC   the new polynucleotides. Specific uses are described for each of the 95
CC   polynucleotides, based on which tissues they are most highly expressed in
CC   (see AAZ24811 for described uses).
XX
SQ   Sequence   114 AA;


  Query Match             70.7%;  Score 41;  DB 20;  Length 114;
  Best Local Similarity   75.0%;  Pred. No. 7.2;
  Matches    6;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 MPFPGPLL 10
              :||||||:
Db         54 lpfpgpli 61


RESULT   3
AAW27100
ID   AAW27100 standard; peptide; 11 AA.
XX
AC   AAW27100;
XX
DT   19-NOV-1997  (first entry)
XX
DE   Angiotensin converting enzyme peptide inhibitor I.
XX
KW   ACE; blood pressure; bradykinin; renal hypertension;
KW   adrenal hypertension; angina pectoris; myocardial infarction;
KW   congestive heart failure; food.
XX
OS   Lactobacillus confusus.
OS   Lactococcus lactis.
XX
PN   JP09188694-A.
XX
PD   22-JUL-1997.
XX
PF   10-JAN-1996;   96JP-0002489.
XX
PR   10-JAN-1996;   96JP-0002489.
XX
PA   (OTSU-) OTSUKA SHOKUHIN KK.
XX
DR   WPI; 1997-420578/39.
XX
PT   Angiotensin converting enzyme peptide inhibitors - useful to
PT   decrease blood pressure and inhibit activation of bradykinin
XX
PS   Claim 1; Page 6; 8pp; Japanese.
XX
CC   The present sequence represents an angiotensin converting enzyme (ACE)
CC   peptide inhibitor. The peptide can be used to decrease blood pressure
CC   and inhibit activation of bradykinin and is useful for the prevention
CC   or treatment of essential, renal or adrenal hypertension. The peptide
CC   is also useful for treatment of increasing the threshold value in
CC   angina pectoris, myocardial infarction, or congestive heart failure.
CC   It is also applicable as an additive to health food or functional foods.
XX
SQ   Sequence   11 AA;
```

```
                     GenCore version 4.5
              Copyright (c) 1993 - 2000  Compugen Ltd.


OM protein - protein search, using sw model

Run on:          October 23, 2001, 13:21:02 ; Search time 46.09 Seconds
                                               (without alignments)
                                               13.153 Million cell updates/sec

Title:           US-09-625-963-3
Perfect score:   58
Sequence:        1 HLMPFPGPLL 10

Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched:        412676 seqs, 60623988 residues

Total number of hits satisfying chosen parameters:      412676

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       A_Geneseq_0601:*
                1:  /SIDS8/gcgdata/geneseq/geneseqp/AA1980.DAT:*
                2:  /SIDS8/gcgdata/geneseq/geneseqp/AA1981.DAT:*
                3:  /SIDS8/gcgdata/geneseq/geneseqp/AA1982.DAT:*
                4:  /SIDS8/gcgdata/geneseq/geneseqp/AA1983.DAT:*
                5:  /SIDS8/gcgdata/geneseq/geneseqp/AA1984.DAT:*
                6:  /SIDS8/gcgdata/geneseq/geneseqp/AA1985.DAT:*
                7:  /SIDS8/gcgdata/geneseq/geneseqp/AA1986.DAT:*
                8:  /SIDS8/gcgdata/geneseq/geneseqp/AA1987.DAT:*
                9:  /SIDS8/gcgdata/geneseq/geneseqp/AA1988.DAT:*
                10:  /SIDS8/gcgdata/geneseq/geneseqp/AA1989.DAT:*
                11:  /SIDS8/gcgdata/geneseq/geneseqp/AA1990.DAT:*
                12:  /SIDS8/gcgdata/geneseq/geneseqp/AA1991.DAT:*
                13:  /SIDS8/gcgdata/geneseq/geneseqp/AA1992.DAT:*
                14:  /SIDS8/gcgdata/geneseq/geneseqp/AA1993.DAT:*
                15:  /SIDS8/gcgdata/geneseq/geneseqp/AA1994.DAT:*
                16:  /SIDS8/gcgdata/geneseq/geneseqp/AA1995.DAT:*
                17:  /SIDS8/gcgdata/geneseq/geneseqp/AA1996.DAT:*
                18:  /SIDS8/gcgdata/geneseq/geneseqp/AA1997.DAT:*
                19:  /SIDS8/gcgdata/geneseq/geneseqp/AA1998.DAT:*
                20:  /SIDS8/gcgdata/geneseq/geneseqp/AA1999.DAT:*
                21:  /SIDS8/gcgdata/geneseq/geneseqp/AA2000.DAT:*
                22:  /SIDS8/gcgdata/geneseq/geneseqp/AA2001.DAT:*

     Pred. No. is the number of results predicted by chance to have a
     score greater than or equal to the score of the result being printed,
     and is derived by analysis of the total score distribution.

                                SUMMARIES

              %
Result      Query
  No.  Score Match Length DB  ID                     Description
------------------------------------------------------------------------
    1     58  100.0     10  21  AAY94204              Human cytotoxic T
    2     41   70.7    114  20  AAY41311              Human secreted pro
    3     40   69.0     11  18  AAW27100              Angiotensin conver
    4     39   67.2      9  15  AAR53617              Opioid peptide #11
    5     39   67.2     10  15  AAR53622              Opioid peptide #16
    6     39   67.2     19  21  AAB26103              Stereoselective tr
    7     38   65.5     50  20  AAY35905              Extended human sec
    8     38   65.5     50  20  AAY13094              Human secreted pro
    9     38   65.5    112  20  AAW78145              Human secreted pro
   10     38   65.5    360  21  AAY77470              Human deubiquitina
   11     38   65.5    498  16  AAR82242              Capsanthin-capsoru
   12     38   65.5    550  20  AAY08899              S. peregrina cyste
   13     37   63.8    778  17  AAR94900              Human tastin.  Hom
   14     37   63.8    778  21  AAB23359              Human tastin prote
   15     37   63.8   1069  21  AAB36839              Protein sequence a
   16     37   63.8   1920  22  AAB65656              Novel protein kina
   17     36   62.1      8  13  AAR26796              Prolyl endopeptida
   18     36   62.1      9  13  AAR26798              Prolyl endopeptida
   19     36   62.1      9  13  AAR26797              Prolyl endopeptida
   20     36   62.1    149  20  AAY74068              Human prostate tum
   21     36   62.1    161  21  AAY87972              Pseudomonas sp. HR
   22     36   62.1    225  22  AAB63952              Human prostate can
   23     36   62.1    336  21  AAY87971              Pseudomonas sp. HR
   24     36   62.1    423  21  AAY87970              Pseudomonas sp. HR
   25     36   62.1    527  20  AAY29515              Human lung tumour
   26     36   62.1    527  21  AAB44492              Human lung tumour-
   27     36   62.1    739  21  AAY77113              Ebola virus nucleo
   28     36   62.1   1089  21  AAY99369              Human PRO1249 (UNQ
   29     36   62.1   1089  22  AAB66118              Protein of the inv
   30     36   62.1   1996  18  AAW22607              Platenolide syntha
   31     36   62.1   1996  18  AAW23717              Platenolide syntha
   32     35   60.3      6  21  AAY49376              Beta-casomorphin d
   33     35   60.3      7   6  AAP50912             Sequence of morphi
   34     35   60.3      7  13  AAR27171              Sequence of peptid
   35     35   60.3      7  15  AAR53416              Peptide having opi
   36     35   60.3      7  15  AAR53623              Opioid peptide #17
   37     35   60.3      7  21  AAY49371              Beta-casomorphin d
   38     35   60.3      7  22  AAB74637              Milk protein fragm
   39     35   60.3      8  15  AAR53626              Opioid peptide #20
   40     35   60.3      9  13  AAR26793              Prolyl endopeptida
   41     35   60.3      9  13  AAR26794              Prolyl endopeptida
   42     35   60.3      9  13  AAR26795              Prolyl endopeptida
   43     35   60.3      9  15  AAR53613              Opioid peptide #7.
   44     35   60.3      9  15  AAR53614              Opioid peptide #8.
   45     35   60.3      9  15  AAR53615              Opioid peptide #9.


                              ALIGNMENTS


RESULT   1
AAY94204
ID   AAY94204 standard; peptide; 10 AA.
XX
AC   AAY94204;
XX
DT   28-JUL-2000  (first entry)
XX
DE   Human cytotoxic T lymphocyte-recognised gata-1 peptide hug 378-87.
XX
KW   Hug 378-87; peptide; epitope; gata-1 gene; leukaemia; immunotherapy;
KW   transcription factor.
XX
OS   Homo sapiens.
XX
PN   WO200026249-A1.
XX
PD   11-MAY-2000.
XX
PF   02-NOV-1999;   99WO-GB03572.
XX
PR   02-NOV-1998;   98GB-0023897.
XX
PA   (IMCO-) IMPERIAL COLLEGE INNOVATIONS LTD.
XX
PI   Stauss HJ,  Gao L;
XX
DR   WPI; 2000-376123/32.
XX
PT   Novel peptides comprising WT-1 and GATA-1 epitopes, their fragments or
PT   variants, useful as vaccines for cancer immunotherapy  -
XX
PS   Claim 3; Page 74; 93pp; English.
XX
```

```
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=OT3;
RX   MEDLINE=98344137; PubMed=9679194;
RA   Kawarabayasi Y., Sawada M., Horikawa H., Haikawa Y., Hino Y.,
RA   Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyama A., Nagai Y.,
RA   Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohfuku Y.,
RA   Funahashi T., Tanaka T., Kudoh Y., Yamazaki J., Kushida N., Oguchi A.,
RA   Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K.,
RA   Masuchi Y., Shizuya H., Kikuchi H.;
RT   "Complete sequence and gene organization of the genome of a hyper-
RT   thermophilic archaebacterium, Pyrococcus horikoshii OT3.";
RL   DNA Res. 5:55-76(1998).
DR   EMBL; AP000006; BAA30453.1; -.
DR   HSSP; P04079; 1GPM.
DR   InterPro; IPR001674; -.
DR   Pfam; PF00958; GMP_synt_C; 1.
SQ   SEQUENCE   308 AA;  34562 MW;  8B209770605AE4BA CRC64;

  Query Match             65.5%;  Score 38;  DB 1;  Length 308;
  Best Local Similarity   100.0%;  Pred. No. 62;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      3 MPFPGP 8
          ||||||
Db    186 MPFPGP 191


Search completed: October 23, 2001, 13:27:55
Job time: 408 sec
```

```
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   ARX HOMEODOMAIN PROTEIN.
GN   ARX.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BALB-C;
RA   Ohsaki K.;
RT   "Homeobox genes and nervous development.";
RL   Thesis (1999), Nara Institute of Science and Technology.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC   -!- SIMILARITY: TO OTHER HOMEOBOX DOMAINS.
DR   EMBL; AB026674; BAA85852.1; -.
DR   EMBL; AB026670; BAA85852.1; JOINED.
DR   EMBL; AB026671; BAA85852.1; JOINED.
DR   EMBL; AB026672; BAA85852.1; JOINED.
DR   EMBL; AB026673; BAA85852.1; JOINED.
DR   HSSP; P02833; 1SAN.
DR   InterPro; IPR000104; -.
DR   InterPro; IPR001356; -.
DR   InterPro; IPR003654; -.
DR   Pfam; PF00046; homeobox; 1.
DR   PRINTS; PR00308; ANTIFREEZEI.
DR   PROSITE; PS00027; HOMEOBOX_1; 1.
DR   PROSITE; PS50071; HOMEOBOX_2; 1.
DR   SMART; SM00389; HOX; 1.
KW   DNA-binding; Homeobox; Nuclear protein.
SQ   SEQUENCE   562 AA;  58215 MW;  A64AEA1C351FC302 CRC64;


  Query Match             67.2%;  Score 39;  DB 11;  Length 562;
  Best Local Similarity   85.7%;  Pred. No. 76;
  Matches    6;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        3 MPFPGPL 9
            :||||||
Db      399 LPFPGPL 405


RESULT  13
O29986
ID   O29986      PRELIMINARY;      PRT;   303 AA.
AC   O29986;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   GMP SYNTHASE (GUAA-1).
GN   AF0253.
OS   Archaeoglobus fulgidus.
OC   Archaea; Euryarchaeota; Archaeoglobales; Archaeoglobaceae;
OC   Archaeoglobus.
OX   NCBI_TaxID=2234;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=VC-16 / DSM 4304 / ATCC 49558;
RX   MEDLINE=98049343; PubMed=9389475;
RA   Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
RA   Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,
RA   Richardson D.L., Kerlavage A.R., Graham D.E., Kyrpides N.C.,
RA   Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
RA   Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
RA   Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
RA   Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,
RA   Cotton M.D., Spriggs T., Artiach P., Kaine B.P., Sykes S.M.,
RA   Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA   Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA   Venter J.C.;
RT   "The complete genome sequence of the hyperthermophilic, sulphate-
```

```
RT   reducing archaeon Archaeoglobus fulgidus.";
RL   Nature 390:364-370(1997).
DR   EMBL; AE001088; AAB90977.1; -.
DR   HSSP; P04079; 1GPM.
DR   TIGR; AF0253; -.
DR   InterPro; IPR001674; -.
DR   Pfam; PF00958; GMP_synt_C; 1.
KW   Hypothetical protein.
SQ   SEQUENCE   303 AA;  34119 MW;  FC6259333E624BC2 CRC64;

  Query Match             65.5%;  Score 38;  DB 1;  Length 303;
  Best Local Similarity   100.0%;  Pred. No. 61;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        3 MPFPGP 8
              ||||||
Db      184 MPFPGP 189


RESULT  14
Q9HP33
ID   Q9HP33      PRELIMINARY;      PRT;   305 AA.
AC   Q9HP33;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   GMP SYNTHASE SUBUNIT B.
GN   GUAAB OR VNG1829G.
OS   Halobacterium sp. (strain NRC-1).
OC   Archaea; Euryarchaeota; Halobacteriales; Halobacteriaceae;
OC   Halobacterium.
OX   NCBI_TaxID=64091;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20504483; PubMed=11016950;
RA   Ng W.V., Kennedy S.P., Mahairas G.G., Berquist B., Pan M.,
RA   Shukla H.D., Lasky S.R., Baliga N.S., Thorsson V., Sbrogna J.,
RA   Swartzell S., Weir D., Hall J., Dahl T.A., Welti R., Goo Y.A.,
RA   Leithauser B., Keller K., Cruz R., Danson M.J., Hough D.W.,
RA   Maddocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H.,
RA   Isenbarger T.A., Peck R.F., Pohlschroder M., Spudich J.L., Jung K.-H.,
RA   Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
RA   Ebhardt H., Lowe T.M., Liang P., Riley M., Hood L., DasSarma S.;
RT   "Genome sequence of Halobacterium species NRC-1.";
RL   Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
DR   EMBL; AE005084; AAG20037.1; -.
DR   InterPro; IPR001674; -.
DR   Pfam; PF00958; GMP_synt_C; 1.
SQ   SEQUENCE   305 AA;  33724 MW;  D47A25946A0869ED CRC64;

  Query Match             65.5%;  Score 38;  DB 1;  Length 305;
  Best Local Similarity   100.0%;  Pred. No. 62;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        3 MPFPGP 8
              ||||||
Db      186 MPFPGP 191


RESULT  15
O59072
ID   O59072      PRELIMINARY;      PRT;   308 AA.
AC   O59072;
DT   01-AUG-1998 (TrEMBLrel. 07, Created)
DT   01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   308AA LONG HYPOTHETICAL GMP SYNTHASE.
GN   PH1347.
OS   Pyrococcus horikoshii.
OC   Archaea; Euryarchaeota; Thermococcales; Thermococcaceae; Pyrococcus.
OX   NCBI_TaxID=53953;
```

```
RESULT   9
Q88303
ID   Q88303      PRELIMINARY;      PRT;   190 AA.
AC   Q88303;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   ENVELOPE PROTEIN GP20E (FRAGMENT).
OS   Simian SRV-like type D retrovirus.
OC   Viruses; Retroid viruses; Retroviridae; Type D retroviruses.
OX   NCBI_TaxID=36771;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PAPIO CYNOCEPHALUS;
RX   MEDLINE=95176553; PubMed=7871741;
RA   Grant R.F., Windsor S.K., Malinak C.J., Bartz C.R., Sabo A.,
RA   Benveniste R.E., Tsai C.;
RT   "Characterization of infectious type D retrovirus from baboons.";
RL   Virology 207:292-296(1995).
DR   EMBL; U16844; AAC54042.1; -.
DR   HSSP; P03385; 1MOF.
DR   InterPro; IPR002050; -.
DR   Pfam; PF00429; ENV_polyprotein; 1.
KW   Envelope protein.
FT   NON_TER       1      1
SQ   SEQUENCE   190 AA;  20975 MW;  F7F2A3DFCBABC770 CRC64;


  Query Match             67.2%;  Score 39;  DB 14;  Length 190;
  Best Local Similarity   70.0%;  Pred. No. 26;
  Matches    7;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 HLMPFPGPLL 10
                :|:|| ||||
Db        132 YLLPFLGPLL 141


RESULT  10
Q9M3W3
ID   Q9M3W3      PRELIMINARY;      PRT;   251 AA.
AC   Q9M3W3;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   PUTATIVE BOWMAN BIRK TRYPSIN INHIBITOR.
GN   RBBI3-1.
OS   Oryza sativa (Rice).
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae;
OC   Oryza.
OX   NCBI_TaxID=4530;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. TEQING; TISSUE=LEAF;
RA   Qu L., Gu H., Liu M., Li D., Zhu G., Li N., Pan N., Chen Z.;
RT   "Molecular cloning, expression pattern and evolution of rice Bowman
RT   Birk trypsin inhibitor gene family.";
RL   Submitted (APR-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AJ277468; CAB88208.1; -.
DR   InterPro; IPR000877; -.
DR   Pfam; PF00228; Bowman-Birk_leg; 2.
DR   PROSITE; PS00281; BOWMAN_BIRK; 2.
DR   SMART; SM00269; BowB; 1.
SQ   SEQUENCE   251 AA;  27785 MW;  3F4C53B2CAB149AD CRC64;


  Query Match             67.2%;  Score 39;  DB 10;  Length 251;
  Best Local Similarity   85.7%;  Pred. No. 35;
  Matches    6;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          4 PFPGPLL 10
                ||||||:
Db        162 PFPGPLI 168


RESULT  11
O42115
ID   O42115      PRELIMINARY;      PRT;   453 AA.
AC   O42115;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   ARISTALESS RELATED HOMEOBOX PROTEIN (ARX).
GN   ARX.
OS   Brachydanio rerio (Zebrafish) (Zebra danio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
OC   Cypriniformes; Cyprinidae; Rasborinae; Danio.
OX   NCBI_TaxID=7955;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=EMBRYO;
RX   MEDLINE=97398450; PubMed=9256348;
RA   Miura H., Yanazawa M., Kato K., Kitamura K.;
RT   "Expression of a novel aristaless related homeobox gene 'Arx' in the
RT   vertebrate telencephalon, diencephalon and floor plate.";
RL   Mech. Dev. 65:99-109(1997).
CC   -!- FUNCTION: APPEARS TO BE INDISPENSABLE FOR THE CENTRAL NERVOUS
CC       SYSTEM DEVELOPMENT. MAY HAVE A ROLE IN THE NEURONAL
CC       DIFFERENTIATION OF THE GANGLIONIC EMINENCE AND VENTRAL THALAMUS.
CC       MAY ALSO BE INVOLVED IN AXONAL GUIDANCE IN THE FLOOR PLATE.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN BRAIN.
CC   -!- DEVELOPMENTAL STAGE: EXPRESSED AT 10H AND 12H IN DIENCEPHALON.
CC       EXPRESSED TRANSIENTLY AT 12H IN CAUDAL TELENCEPHALON. LATER
CC       EXPRESSION IN FLOOR PLATE AND SOMITES, FOLLOWED BY ROSTRAL
CC       TELENCEPHALON AND VENTRAL THALAMUS. EXPRESSED AT 40H IN
CC       HYPOTHALAMUS.
CC   -!- DOMAIN: A 14 AMINO ACID MOTIF (OAR-DOMAIN) IS CONSERVED AMONG
CC       SEVERAL HD PROTEINS. MAY BE INVOLVED IN ADDITIONAL DNA-BINDING
CC       (OUTSIDE THE HD-DNA INTERACTION) OR MAY BE A SITE OF PROTEIN-
CC       PROTEIN INTERACTION. IT COULD HAVE A ROLE IN TRANSACTIVATION.
CC   -!- SIMILARITY: BELONGS TO THE PAIRED FAMILY OF HOMEOBOX PROTEINS.
DR   EMBL; AB006104; BAA21764.1; -.
DR   HSSP; P06601; 1FJL.
DR   ZFIN; ZDB-GENE-990415-15; arx.
DR   InterPro; IPR001356; -.
DR   InterPro; IPR003654; -.
DR   Pfam; PF00046; homeobox; 1.
DR   PRINTS; PR00024; HOMEOBOX.
DR   PROSITE; PS00027; HOMEOBOX_1; 1.
DR   PROSITE; PS50071; HOMEOBOX_2; 1.
DR   SMART; SM00389; HOX; 1.
KW   Homeobox; DNA-binding; Developmental protein; Nuclear protein;
KW   Transcription regulation; Activator.
FT   DOMAIN      320    327       POLY-ALA.
FT   DOMAIN      409    413       POLY-SER.
FT   DNA_BIND    215    274       HOMEOBOX.
FT   DOMAIN      421    434       OAR DOMAIN.
SQ   SEQUENCE   453 AA;  49396 MW;  547F7CC478534808 CRC64;


  Query Match             67.2%;  Score 39;  DB 13;  Length 453;
  Best Local Similarity   85.7%;  Pred. No. 61;
  Matches    6;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 MPFPGPL 9
                :||||||
Db        285 LPFPGPL 291


RESULT  12
Q9QYT4
ID   Q9QYT4      PRELIMINARY;      PRT;   562 AA.
AC   Q9QYT4;
```

```
RT   "Sequencing the distal X chromosome of Drosophila melanogaster.";
RL   Submitted (DEC-1997) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Benos P.;
RL   Submitted (APR-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AL009188; CAA15674.2; -.
DR   FlyBase; FBgn0001337; l(1)1Bb.
DR   InterPro; IPR001494; -.
SQ   SEQUENCE   1060 AA;  119276 MW;  650B03CD25DB9156 CRC64;

  Query Match             70.7%;  Score 41;  DB 5;  Length 1060;
  Best Local Similarity   75.0%;  Pred. No. 64;
  Matches    6;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        3 MPFPGPLL 10
            :||||||:
Db      302 LPFPGPLM 309


RESULT   6
O77031
ID   O77031      PRELIMINARY;      PRT;   346 AA.
AC   O77031;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   ACHAETE-SCUTE COMPLEX PROTEIN SC (SCUTE PROTEIN).
GN   SC.
OS   Drosophila simulans (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7240;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SIM-5 G20;
RX   MEDLINE=98278813; PubMed=9611206;
RA   Takano T.S.;
RT   "Rate variation of DNA sequence evolution in the Drosophila
RT   lineages.";
RL   Genetics 149:959-970(1998).
CC   -!- FUNCTION: AS-C PROTEINS ARE INVOLVED IN THE DETERMINATION OF THE
CC       NEURONAL PRECURSORS IN THE PERIPHERAL NERVOUS SYSTEM AND THE
CC       CENTRAL NERVOUS SYSTEM. ALSO INVOLVED IN SEX DETERMINATION AND
CC       DOSAGE COMPENSATION (BY SIMILARITY).
CC   -!- SUBUNIT: EFFICIENT DNA BINDING REQUIRES DIMERIZATION WITH ANOTHER
CC       BHLH PROTEIN (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF
CC       TRANSCRIPTION FACTORS.
DR   EMBL; AB005801; BAA33212.1; -.
DR   FlyBase; FBgn0012893; Dsim\sc.
DR   InterPro; IPR001092; -.
DR   InterPro; IPR003015; -.
DR   Pfam; PF00010; HLH; 1.
DR   PROSITE; PS00038; HELIX_LOOP_HELIX; 1.
DR   SMART; SM00353; HLH; 1.
KW   Neurogenesis; Differentiation; Developmental protein.
FT   DNA_BIND    102    112       BASIC DOMAIN (BY SIMILARITY).
FT   DOMAIN      113    163       HELIX-LOOP-HELIX MOTIF (BY SIMILARITY).
FT   SIMILAR     332    346       TO AC AND SC PROTEINS.
SQ   SEQUENCE   346 AA;  38321 MW;  1433B75DBC0A534A CRC64;

  Query Match             69.0%;  Score 40;  DB 5;  Length 346;
  Best Local Similarity   60.0%;  Pred. No. 32;
  Matches    6;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy        1 HLMPFPGPLL 10
            |:|| | ||:
Db       39 HIMPAPSPLI 48
```

```
RESULT   7
Q53787
ID   Q53787      PRELIMINARY;      PRT;   441 AA.
AC   Q53787;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   ORGANIC ACID TRANSPORT PROTEIN.
OS   Streptococcus bovis.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC   Streptococcus.
OX   NCBI_TaxID=1315;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 15352;
RX   MEDLINE=96316385; PubMed=8702261;
RA   Kawai S., Suzuki H., Yamamoto K., Inui M., Yukawa H., Kumagai H.;
RT   "Purification and characterization of a malic enzyme from the ruminal
RT   bacterium Streptococcus bovis ATCC 15352 and cloning and sequencing of
RT   its gene.";
RL   Appl. Environ. Microbiol. 62:2692-2700(1996).
DR   EMBL; U35658; AAB18291.1; -.
SQ   SEQUENCE   441 AA;  47221 MW;  189F1C2125E5C559 CRC64;

  Query Match             69.0%;  Score 40;  DB 2;  Length 441;
  Best Local Similarity   70.0%;  Pred. No. 40;
  Matches    7;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy        1 HLMPFPGPLL 10
            ||  ||||:|
Db      290 HLTGFPGPVL 299


RESULT   8
Q9M3W1
ID   Q9M3W1      PRELIMINARY;      PRT;   185 AA.
AC   Q9M3W1;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   PUTATIVE BOWMAN BIRK TRYPSIN INHIBITOR.
GN   RBBI2-3.
OS   Oryza sativa (Rice).
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae;
OC   Oryza.
OX   NCBI_TaxID=4530;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. TEQING;
RA   Qu L., Gu H., Liu M., Li D., Zhu G., Li N., Pan N., Chen Z.;
RT   "Molecular cloning, expression pattern and evolution of rice Bowman
RT   Birk trypsin inhibitor family.";
RL   Submitted (APR-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AJ277470; CAB88391.1; -.
DR   InterPro; IPR000877; -.
DR   Pfam; PF00228; Bowman-Birk_leg; 1.
DR   ProDom; PD002168; -; 2.
DR   PROSITE; PS00281; BOWMAN_BIRK; 1.
DR   SMART; SM00269; BowB; 1.
SQ   SEQUENCE   185 AA;  20089 MW;  49E72E1F62490345 CRC64;

  Query Match             67.2%;  Score 39;  DB 10;  Length 185;
  Best Local Similarity   85.7%;  Pred. No. 26;
  Matches    6;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        4 PFPGPLL 10
            ||||||:
Db       96 PFPGPLI 102
```

```
  Query Match             72.4%;  Score 42;  DB 5;  Length 336;
  Best Local Similarity   70.0%;  Pred. No. 14;
  Matches    7;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 HLMPFPGPLL 10
          |:|| | |||
Db     27 HIMPAPSPLL 36


RESULT   2
Q9I6J9
ID   Q9I6J9      PRELIMINARY;      PRT;   368 AA.
AC   Q9I6J9;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   HYPOTHETICAL PROTEIN PA0292.
GN   PA0292.
OS   Pseudomonas aeruginosa.
OC   Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC   Pseudomonas.
OX   NCBI_TaxID=287;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PAO1;
RX   MEDLINE=20437337; PubMed=10984043;
RA   Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warrener P.,
RA   Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA   Garber R.L., Goltry L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
RA   Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA   Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA   Reizer J., Saier M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT   "Complete genome sequence of Pseudomonas aeruginosa PA01, an
RT   opportunistic pathogen.";
RL   Nature 406:959-964(2000).
DR   EMBL; AE004467; AAG03681.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   368 AA;  41191 MW;  E805F3C907AC1E79 CRC64;


  Query Match             72.4%;  Score 42;  DB 2;  Length 368;
  Best Local Similarity   77.8%;  Pred. No. 16;
  Matches    7;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 HLMPFPGPL 9
          | || ||||
Db    267 HKMPIPGPL 275


RESULT   3
Q9RC03
ID   Q9RC03      PRELIMINARY;      PRT;   376 AA.
AC   Q9RC03;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   HYPOTHETICAL 41.0 KDA PROTEIN.
OS   Pseudomonas sp. BG33R.
OC   Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC   Pseudomonas.
OX   NCBI_TaxID=96901;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BG33R;
RA   Glandorf D.C.M., Wechter W.P., Derrick W.C., Leverentz B.,
RA   Kluepfel D.A.;
RT   "Identification of genetic loci in a rhizosphere inhabiting
RT   fluorescent Pseudomonas sp. BG33R involved in expression of a
RT   phytoparasitic nematode ovicidal factor.";
RL   Submitted (MAY-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF153708; AAD47622.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   376 AA;  41026 MW;  94D6A0D859F39AFE CRC64;

  Query Match             72.4%;  Score 42;  DB 2;  Length 376;
  Best Local Similarity   77.8%;  Pred. No. 16;
  Matches    7;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 HLMPFPGPL 9
          | || ||||
Db    267 HKMPIPGPL 275


RESULT   4
Q9LL57
ID   Q9LL57      PRELIMINARY;      PRT;   650 AA.
AC   Q9LL57;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   RECEPTOR-LIKE KINASE.
GN   RLG15.
OS   Oryza sativa (Rice).
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae;
OC   Oryza.
OX   NCBI_TaxID=4530;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Yun C.-H., Lee G.-R., Kim H.-I.;
RT   "Molecular analysis of receptor-like kinase in rice.";
RL   Submitted (FEB-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR   EMBL; AF238472; AAF78016.1; -.
DR   InterPro; IPR000719; -.
DR   InterPro; IPR001917; -.
DR   InterPro; IPR002290; -.
DR   Pfam; PF00069; pkinase; 1.
DR   PROSITE; PS00599; AA_TRANSFER_CLASS_2; UNKNOWN_1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR   SMART; SM00220; S_TKc; 1.
KW   ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ   SEQUENCE   650 AA;  73103 MW;  608B6771A9347B3F CRC64;


  Query Match             72.4%;  Score 42;  DB 10;  Length 650;
  Best Local Similarity   77.8%;  Pred. No. 27;
  Matches    7;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 HLMPFPGPL 9
          ||:||| ||
Db    106 HLLPFPCPL 114


RESULT   5
O46055
ID   O46055      PRELIMINARY;      PRT;  1060 AA.
AC   O46055;
DT   01-JUN-1998 (TrEMBLrel. 06, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   EG:165H7.3 PROTEIN.
GN   L(1)1BB OR EG:165H7.3 OR CG3923.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   de Pablos B., Madueno E., Modolell J.;
```

```
                       GenCore version 4.5
              Copyright (c) 1993 - 2000  Compugen Ltd.

OM protein - protein search, using sw model

Run on:         October 23, 2001, 13:27:54 ; Search time 47.11 Seconds
                                             (without alignments)
                                             28.084 Million cell updates/sec

Title:          US-09-625-963-3
Perfect score:  58
Sequence:       1 HLMPFPGPLL 10

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched:       425026 seqs, 132305027 residues

Total number of hits satisfying chosen parameters:      425026

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :      SPTREMBL_16:*
               1:  sp_archea:*
               2:  sp_bacteria:*
               3:  sp_fungi:*
               4:  sp_human:*
               5:  sp_invertebrate:*
               6:  sp_mammal:*
               7:  sp_mhc:*
               8:  sp_organelle:*
               9:  sp_phage:*
               10:  sp_plant:*
               11:  sp_rodent:*
               12:  sp_unclassified:*
               13:  sp_vertebrate:*
               14:  sp_virus:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                                  SUMMARIES
                  %
Result          Query
  No.    Score  Match Length DB  ID                          Description
 -----------------------------------------------------------------------------
     1      42   72.4    336  5  O77029                      O77029 drosophila
     2      42   72.4    368  2  Q9I6J9                      Q9i6j9 pseudomonas
     3      42   72.4    376  2  Q9RC03                      Q9rc03 pseudomonas
     4      42   72.4    650  10  Q9LL57                      Q91157 oryza sativ
     5      41   70.7   1060  5  O46055                      O46055 drosophila
     6      40   69.0    346  5  O77031                      O77031 drosophila
     7      40   69.0    441  2  Q53787                      Q53787 streptococc
     8      39   67.2    185  10  Q9M3W1                      Q9m3w1 oryza sativ
     9      39   67.2    190  14  Q88303                      Q88303 simian srv-
    10      39   67.2    251  10  Q9M3W3                      Q9m3w3 oryza sativ
    11      39   67.2    453  13  O42115                      O42115 brachydanio
    12      39   67.2    562  11  Q9QYT4                      Q9qyt4 mus musculu
    13      38   65.5    303  1  O29986                      O29986 archaeoglob
    14      38   65.5    305  1  Q9HP33                      Q9hp33 halobacteri
    15      38   65.5    308  1  O59072                      O59072 pyrococcus
    16      38   65.5    308  1  Q9V0I7                      Q9v0i7 pyrococcus
    17      38   65.5    399  2  Q9HW00                      Q9hw00 pseudomonas
    18      38   65.5    550  5  O97453                      O97453 sarcophaga
    19      38   65.5    723  4  Q9H899                      Q9h899 homo sapien
    20      38   65.5    753  10  Q9SDD2                      Q9sdd2 oryza sativ
    21      38   65.5    850  11  Q9JJ15                      Q9jj15 mus musculu
    22      37   63.8     68  5  Q9NME9                      Q9nme9 leishmania
    23      37   63.8    116  2  P72878                      P72878 synechocyst
    24      37   63.8    154  14  Q9YKJ9                      Q9ykj9 hepatitis b
    25      37   63.8    154  14  Q9IF29                      Q9if29 hepatitis b
    26      37   63.8    310  1  Q9HIH8                      Q9hih8 thermoplasm
    27      37   63.8    420  2  Q9RJQ6                      Q9rjq6 streptomyce
    28      37   63.8    833  2  Q9RZT0                      Q9rzt0 deinococcus
    29      37   63.8   1246  4  O15052                      O15052 homo sapien
    30      37   63.8   2126  11  Q9JIH7                      Q9jih7 rattus norv
    31      37   63.8   2382  4  Q9H4A3                      Q9h4a3 homo sapien
    32      36   62.1     69  2  Q9EWV8                      Q9ewv8 streptomyce
    33      36   62.1    145  4  Q9P1S1                      Q9p1s1 homo sapien
    34      36   62.1    246  2  O83043                      O83043 azospirillu
    35      36   62.1    275  5  Q9VP40                      Q9vp40 drosophila
    36      36   62.1    343  5  Q22807                      Q22807 caenorhabdi
    37      36   62.1    478  11  Q9WUF2                      Q9wuf2 mus musculu
    38      36   62.1    589  2  Q9RLD9                      Q9rld9 pseudomonas
    39      36   62.1    739  14  Q9DQD3                      Q9dqd3 ebola virus
    40      36   62.1    954  5  Q22360                      Q22360 caenorhabdi
    41      36   62.1   1000  5  Q9U3H5                      Q9u3h5 caenorhabdi
    42      36   62.1   1159  11  Q9EPR5                      Q9epr5 mus musculu
    43    35.5   61.2   1054  10  Q9FKL3                      Q9fkl3 arabidopsis
    44      35   60.3    100  2  O87245                      O87245 lactococcus
    45      35   60.3    102  1  O59168                      O59168 pyrococcus

                                 ALIGNMENTS

RESULT   1
O77029
ID   O77029      PRELIMINARY;      PRT;   336 AA.
AC   O77029;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   ACHAETE-SCUTE COMPLEX PROTEIN SC (SCUTE PROTEIN).
GN   SC.
OS   Drosophila yakuba (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7245;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=IVORY COAST;
RX   MEDLINE=98278813; PubMed=9611206;
RA   Takano T.S.;
RT   "Rate variation of DNA sequence evolution in the Drosophila
RT   lineages.";
RL   Genetics 149:959-970(1998).
CC   -!- FUNCTION: AS-C PROTEINS ARE INVOLVED IN THE DETERMINATION OF THE
CC       NEURONAL PRECURSORS IN THE PERIPHERAL NERVOUS SYSTEM AND THE
CC       CENTRAL NERVOUS SYSTEM. ALSO INVOLVED IN SEX DETERMINATION AND
CC       DOSAGE COMPENSATION (BY SIMILARITY).
CC   -!- SUBUNIT: EFFICIENT DNA BINDING REQUIRES DIMERIZATION WITH ANOTHER
CC       BHLH PROTEIN (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF
CC       TRANSCRIPTION FACTORS.
DR   EMBL; AB005799; BAA33210.1; -.
DR   FlyBase; FBgn0025397; Dyak\sc.
DR   InterPro; IPR001092; -.
DR   InterPro; IPR003015; -.
DR   Pfam; PF00010; HLH; 1.
DR   PROSITE; PS00038; HELIX_LOOP_HELIX; 1.
DR   SMART; SM00353; HLH; 1.
KW   Neurogenesis; Differentiation; Developmental protein.
FT   DNA_BIND      90    100       BASIC DOMAIN (BY SIMILARITY).
FT   DOMAIN       101    151       HELIX-LOOP-HELIX MOTIF (BY SIMILARITY).
FT   SIMILAR      322    336       TO AC AND L(1)SC PROTEINS.
SQ   SEQUENCE   336 AA;  37050 MW;  0202BB37BCB1A9BC CRC64;
```

```
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; Y09358; CAA70541.1; -.
KW   Nucleoprotein.
SQ   SEQUENCE   739 AA;  83240 MW;  33A96720FFBC638E CRC64;


  Query Match             62.1%;  Score 36;  DB 1;  Length 739;
  Best Local Similarity   71.4%;  Pred. No. 86;
  Matches    5;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      3 MPFPGPL 9
          :|||||:
Db    427 IPFPGPI 433


RESULT  14
VNUC_EBOZ5
ID   VNUC_EBOZ5     STANDARD;      PRT;   739 AA.
AC   O72142;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   NUCLEOPROTEIN (NUCLEOCAPSID PROTEIN).
GN   NP.
OS   Ebola virus (strain Zaire-95) (Ebo).
OC   Viruses; ssRNA negative-strand viruses; Mononegavirales; Filoviridae;
OC   Filovirus.
OX   NCBI_TaxID=128951;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98321157; PubMed=9657001;
RA   Vanderzanden L., Bray M., Fuller D., Roberts T., Custer D., Spik K.,
RA   Jahrling P., Huggins J., Schmaljohn A., Schmaljohn C.;
RT   "DNA vaccines expressing either the GP or NP genes of Ebola virus
RT   protect mice from lethal challenge.";
RL   Virology 246:134-144(1998).
CC   -!- FUNCTION: RESPONSIBLE FOR ENCAPSIDATION OF GENOMIC RNA.
CC   -!- DOMAIN: THIS PROTEIN CAN BE DIVIDED INTO A HYDROPHOBIC N-TERMINAL
CC       HALF, AND A HYDROPHILIC AND HIGHLY ACIDIC C-TERMINAL HALF.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AF054908; AAC09342.1; -.
KW   Nucleoprotein.
SQ   SEQUENCE   739 AA;  83316 MW;  74D9437293AFF443 CRC64;


  Query Match             62.1%;  Score 36;  DB 1;  Length 739;
  Best Local Similarity   71.4%;  Pred. No. 86;
  Matches    5;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      3 MPFPGPL 9
          :|||||:
Db    427 IPFPGPI 433


RESULT  15
VNUC_EBOZM
ID   VNUC_EBOZM     STANDARD;      PRT;   739 AA.
AC   P18272; Q9YMG4;
DT   01-NOV-1990 (Rel. 16, Created)
DT   15-JUL-1999 (Rel. 38, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   NUCLEOPROTEIN (NUCLEOCAPSID PROTEIN).
GN   NP.
```

```
OS   Ebola virus (strain Zaire Mayinga) (Ebo).
OC   Viruses; ssRNA negative-strand viruses; Mononegavirales; Filoviridae;
OC   Filovirus.
OX   NCBI_TaxID=128952;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=89243211; PubMed=2718390;
RA   Sanchez A., Kiley M.P., Holloway B.P., McCormick J.B., Auperin D.D.;
RT   "The nucleoprotein gene of Ebola virus: cloning, sequencing, and in
RT   vitro expression.";
RL   Virology 170:81-91(1989).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94055391; PubMed=8237108;
RA   Sanchez A., Kiley M.P., Holloway B.P., Auperin D.D.;
RT   "Sequence analysis of the Ebola virus genome: organization, genetic
RT   elements, and comparison with the genome of Marburg virus.";
RL   Virus Res. 29:215-240(1993).
RN   [3]
RP   REVISIONS.
RA   Sanchez A.;
RL   Submitted (OCT-1997) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RA   Volchkov V.E.;
RL   Submitted (AUG-1998) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: RESPONSIBLE FOR ENCAPSIDATION OF GENOMIC RNA.
CC   -!- DOMAIN: THIS PROTEIN CAN BE DIVIDED INTO A HYDROPHOBIC N-TERMINAL
CC       HALF, AND A HYDROPHILIC AND HIGHLY ACIDIC C-TERMINAL HALF.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; J04337; AAA42977.1; ALT_SEQ.
DR   EMBL; L11365; AAB81001.1; -.
DR   EMBL; AF086833; AAD14590.1; -.
DR   PIR; A31471; VHIWEB.
KW   Nucleoprotein.
SQ   SEQUENCE   739 AA;  83286 MW;  159C254EA0478886 CRC64;


  Query Match             62.1%;  Score 36;  DB 1;  Length 739;
  Best Local Similarity   71.4%;  Pred. No. 86;
  Matches    5;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      3 MPFPGPL 9
          :|||||:
Db    427 IPFPGPI 433


Search completed: October 23, 2001, 13:28:19
Job time: 402 sec
```

```
RESULT  11
EGR4_HUMAN
ID   EGR4_HUMAN     STANDARD;      PRT;   486 AA.
AC   Q05215;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   EARLY GROWTH RESPONSE PROTEIN 4 (EGR-4) (AT133).
GN   EGR4.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93278383; PubMed=8504297;
RA   Holst C., Skerka C., Lichter P., Bialonski A., Zipfel P.F.;
RT   "Genomic organization, chromosomal localization and promoter function
RT   of the human zinc-finger gene pAT133.";
RL   Hum. Mol. Genet. 2:367-372(1993).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92052214; PubMed=1658795;
RA   Mueller H.-J., Skerka C., Bialonski A., Zipfel P.F.;
RT   "Clone pAT133 identifies a gene that encodes another human member of
RT   a class of growth factor-induced genes with almost identical zinc-
RT   finger domains.";
RL   Proc. Natl. Acad. Sci. U.S.A. 88:10079-10083(1991).
CC   -!- SUBCELLULAR LOCATION: NUCLEAR (PROBABLE).
CC   -!- INDUCTION: BY PHA/PMA OR BY SERUM.
CC   -!- SIMILARITY: BELONGS TO THE EGR FAMILY OF C2H2-TYPE ZINC-FINGER
CC       PROTEINS.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; X69438; CAA49214.1; -.
DR   EMBL; X60104; CAA42698.1; -.
DR   PIR; S29992; S29992.
DR   HSSP; P08046; 1A1L.
DR   MIM; 128992; -.
DR   InterPro; IPR000822; -.
DR   Pfam; PF00096; zf-C2H2; 3.
DR   PRINTS; PR00048; ZINCFINGER.
DR   PROSITE; PS00028; ZINC_FINGER_C2H2_1; 3.
DR   PROSITE; PS50157; ZINC_FINGER_C2H2_2; 3.
KW   Nuclear protein; Transcription regulation; DNA-binding; Zinc-finger;
KW   Metal-binding; Repeat.
FT   DOMAIN       380    464       ZINC FINGERS.
FT   ZN_FING      380    407       C2H2-TYPE.
FT   ZN_FING      410    432       C2H2-TYPE.
FT   ZN_FING      438    460       C2H2-TYPE.
FT   CONFLICT     427    427       S -> T (IN REF. 2).
SQ   SEQUENCE   486 AA;  50855 MW;  0DF764427E0A21E3 CRC64;

  Query Match             62.1%;  Score 36;  DB 1;  Length 486;
  Best Local Similarity   85.7%;  Pred. No. 57;
  Matches    6;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      2 LMPFPGP 8
          | |||||
Db    111 LAPFPGP 117

RESULT  12
```

```
VNUC_EBOSB
ID   VNUC_EBOSB     STANDARD;      PRT;   738 AA.
AC   Q9QP77;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   NUCLEOPROTEIN (NUCLEOCAPSID PROTEIN).
GN   NP.
OS   Ebola virus (strain Sudan Boniface) (Ebo).
OC   Viruses; ssRNA negative-strand viruses; Mononegavirales; Filoviridae;
OC   Filovirus.
OX   NCBI_TaxID=128948;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Crise B., Smith J.F., Bray M.;
RT   "Ebola Sudan nucleocapsid protein.";
RL   Submitted (JUL-1999) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: RESPONSIBLE FOR ENCAPSIDATION OF GENOMIC RNA.
CC   -!- DOMAIN: THIS PROTEIN CAN BE DIVIDED INTO A HYDROPHOBIC N-TERMINAL
CC       HALF, AND A HYDROPHILIC AND HIGHLY ACIDIC C-TERMINAL HALF.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AF173836; AAD51107.1; -.
KW   Nucleoprotein.
SQ   SEQUENCE   738 AA;  82008 MW;  21AFE1423CFBF9CD CRC64;

  Query Match             62.1%;  Score 36;  DB 1;  Length 738;
  Best Local Similarity   71.4%;  Pred. No. 86;
  Matches    5;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      3 MPFPGPL 9
          :|||||:
Db    427 IPFPGPI 433

RESULT  13
VNUC_EBOG4
ID   VNUC_EBOG4     STANDARD;      PRT;   739 AA.
AC   Q9QCE9;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   NUCLEOPROTEIN (NUCLEOCAPSID PROTEIN).
GN   NP.
OS   Ebola virus (strain Gabon-94) (Ebo).
OC   Viruses; ssRNA negative-strand viruses; Mononegavirales; Filoviridae;
OC   Filovirus.
OX   NCBI_TaxID=128947;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99036017; PubMed=9820131;
RA   Prehaud C.J.C., Hellebrand E., Coudrier D., Volchkov V.E.,
RA   Volchkova V.A., Feldmann B., Le Guenno B., Bouloy M.;
RT   "Recombinant Ebola virus nucleoprotein and glycoprotein (Gabon 94
RT   strain) provide new tools for the detection of human infections.";
RL   J. Gen. Virol. 79:2565-2572(1998).
CC   -!- FUNCTION: RESPONSIBLE FOR ENCAPSIDATION OF GENOMIC RNA.
CC   -!- DOMAIN: THIS PROTEIN CAN BE DIVIDED INTO A HYDROPHOBIC N-TERMINAL
CC       HALF, AND A HYDROPHILIC AND HIGHLY ACIDIC C-TERMINAL HALF.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
```

```
OC   Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX   NCBI_TaxID=10415;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=89010694; PubMed=3171552;
RA   Okamoto H., Tsuda F., Sakugawa H., Sastrosoewignjo R.I., Imai M.,
RA   Miyakawa Y., Mayumi M.;
RT   "Typing hepatitis B virus by homology in nucleotide sequence:
RT   comparison of surface antigen subtypes.";
RL   J. Gen. Virol. 69:2575-2583(1988).
CC   -!- FUNCTION: INDISPENSABLE FOR HBV REPLICATION IN VIVO.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; D00330; -; NOT_ANNOTATED_CDS.
DR   PIR; JS0603; JS0603.
DR   InterPro; IPR000236; -.
DR   Pfam; PF00739; X; 1.
SQ   SEQUENCE   154 AA;  16558 MW;  E2646C059A30125F CRC64;


  Query Match             63.8%;  Score 37;  DB 1;  Length 154;
  Best Local Similarity   100.0%;  Pred. No. 12;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       4 PFPGPL 9
           ||||||
Db      29 PFPGPL 34


RESULT   9
TAST_HUMAN
ID   TAST_HUMAN     STANDARD;      PRT;   778 AA.
AC   Q12815;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   TROPHININ-ASSOCIATED PROTEIN (TASTIN) (TROPHININ-ASSISTING PROTEIN).
GN   TROAP.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=95278733; PubMed=7758945;
RA   Fukuda M.N., Sato T., Nakayama J., Klier G., Mikami M., Aoki D.,
RA   Nozawa S.;
RT   "Trophinin and tastin, a novel cell adhesion molecule complex with
RT   potential involvement in embryo implantation.";
RL   Genes Dev. 9:1199-1210(1995).
CC   -!- FUNCTION: COULD BE INVOLVED WITH BYSTIN AND TROPHININ IN A CELL
CC       ADHESION MOLECULE COMPLEX THAT MEDIATES AN INITIAL ATTACHMENT OF
CC       THE BLASTOCYST TO UTERINE EPITHELIAL CELLS AT THE TIME OF THE
CC       EMBRYO IMPLANTATION.
CC   -!- SUBUNIT: DIRECTLY BINDS BYSTIN, AND INDIRECTLY TROPHININ.
CC   -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC   -!- TISSUE SPECIFICITY: STRONG EXPRESSION AT IMPLANTATION SITES. WAS
CC       EXCLUSIVELY LOCALIZED TO THE APICAL SIDE OF THE
CC       SYNCYTIOTROPHOBLAST. ALSO FOUND IN MACROPHAGES.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; U04810; AAA79333.1; -.
DR   MIM; 603872; -.
KW   Cell adhesion; Repeat.
FT   DOMAIN      504    687       CYS-RICH.
FT   DOMAIN      516    647       4 X 33 AA APPROXIMATE TANDEM REPEATS.
FT   REPEAT      516    548       1.
FT   REPEAT      549    581       2.
FT   REPEAT      582    614       3.
FT   REPEAT      615    647       4.
SQ   SEQUENCE   778 AA;  83758 MW;  52BAB17165672AC0 CRC64;


  Query Match             63.8%;  Score 37;  DB 1;  Length 778;
  Best Local Similarity   66.7%;  Pred. No. 61;
  Matches    6;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 HLMPFPGPL 9
           |||| | |:
Db     303 HLMPSPAPV 311


RESULT  10
NDPP_MOUSE
ID   NDPP_MOUSE     STANDARD;      PRT;   389 AA.
AC   Q03173;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   01-OCT-1996 (Rel. 34, Last annotation update)
DE   NPC DERIVED PROLINE RICH PROTEIN 1 (NDPP-1).
GN   NDPP1.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RX   MEDLINE=93041923; PubMed=1420303;
RA   Sazuka T., Tomooka Y., Kathju S., Ikawa Y., Noda M., Kumar S.;
RT   "Identification of a developmentally regulated gene in the mouse
RT   central nervous system which encodes a novel proline rich protein.";
RL   Biochim. Biophys. Acta 1132:240-248(1992).
CC   -!- FUNCTION: MAY PLAY AN IMPORTANT BIOLOGICAL ROLE IN THE
CC       DEVELOPMENT AND DIFFERENTIATION OF THE CENTRAL NERVOUS SYSTEM.
CC   -!- TISSUE SPECIFICITY: IN THE HEART AND TESTIS AND LESS SO IN THE
CC       LUNG, SKELETAL MUSCLE, KIDNEY, PANCREAS AND BRAIN.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; D10727; BAA01570.1; -.
DR   MGD; MGI:97291; Ndpp1.
KW   Developmental protein.
FT   DOMAIN       30     52       POLY-PRO.
FT   DOMAIN      163    195       LEU/PRO-RICH.
FT   DOMAIN      273    276       POLY-ARG.
SQ   SEQUENCE   389 AA;  41249 MW;  EE36C1CBE8156033 CRC64;


  Query Match             62.1%;  Score 36;  DB 1;  Length 389;
  Best Local Similarity   60.0%;  Pred. No. 46;
  Matches    6;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 HLMPFPGPLL 10
           ||:| | | |
Db     186 HLLPLPSPHL 195
```

```
GN   ENV.
OS   Baboon endogenous virus (strain M7).
OC   Viruses; Retroid viruses; Retroviridae, Mammalian type C retroviruses.
OX   NCBI_TaxID=11764;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Kato S., Matsuo K., Nishimura N., Takahashi N., Takano T.;
RT   "The entire nucleotide sequence of baboon endogenous virus DNA: a
RT   chimeric genome structure of murine type C and simian type D
RT   retroviruses.";
RL   Jpn. J. Genet. 62:127-137(1987).
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; D10032; BAA00924.1; -.
DR   EMBL; X05470; CAA29028.1; -.
DR   PIR; JT0262; VCMVM7.
DR   HSSP; P03385; 1MOF.
DR   InterPro; IPR002050; -.
DR   Pfam; PF00429; ENV_polyprotein; 1.
KW   Coat protein; Glycoprotein; Polyprotein; Transmembrane; Signal.
FT   SIGNAL        1     20       POTENTIAL.
FT   CHAIN        21    346       OUTER MEMBRANE PROTEIN GP70 (POTENTIAL).
FT   CHAIN       347    563       TRANSMEMBRANE PROTEIN P20E (POTENTIAL).
FT   CARBOHYD    113    113       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    219    219       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    229    229       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    264    264       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    282    282       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    292    292       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    306    306       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    312    312       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    321    321       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    339    339       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    469    469       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   563 AA;  61879 MW;  9573137DC4620BB7 CRC64;


  Query Match             67.2%;  Score 39;  DB 1;  Length 563;
  Best Local Similarity   70.0%;  Pred. No. 21;
  Matches    7;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 HLMPFPGPLL 10
           :|:|| ||||
Db     507 YLLPFLGPLL 516


RESULT   7
CCS_CAPAN
ID   CCS_CAPAN      STANDARD;      PRT;   498 AA.
AC   Q42435; Q39470;
DT   01-OCT-2000 (Rel. 40, Created)
DT   01-OCT-2000 (Rel. 40, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   CAPSANTHIN/CAPSORUBIN SYNTHASE, CHLOROPLAST PRECURSOR.
GN   CCS.
OS   Capsicum annuum (Bell pepper).
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I;
OC   Solanales; Solanaceae; Capsicum.
OX   NCBI_TaxID=4072;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. YOLO WONDER;
RX   MEDLINE=94197697; PubMed=8147854;
RA   Deruere J., Bouvier F., Steppuhn J., Klein A., Camara B., Kuntz M.;
RT   "Structure and expression of two plant genes encoding chromoplast-
RT   specific proteins: occurrence of partially spliced transcripts.";
RL   Biochem. Biophys. Res. Commun. 199:1144-1150(1994).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. LAMUYO;
RX   MEDLINE=95004653; PubMed=7920703;
RA   Bouvier F., Huqueney P., d'Harlinque A., Kuntz M., Camara B.;
RT   "Xanthophyll biosynthesis in chromoplasts: isolation and molecular
RT   cloning of an enzyme catalyzing the conversion of 5,6-epoxycarotenoid
RT   into ketocarotenoid.";
RL   Plant J. 6:45-54(1994).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. YOLO WONDER;
RX   MEDLINE=95179811; PubMed=7874747;
RA   Houln G., Schantz M.L., Meyer B., Pozueta-Romero J., Schantz R.;
RT   "A chromoplast-specific protein in Capsicum annuum: characterization
RT   and expression of the corresponding gene.";
RL   Curr. Genet. 26:524-527(1994).
CC   -!- FUNCTION: CATALYZES THE CONVERSION OF THE UBIQUITOUS 5,6-
CC       EPOXYCAROTENOIDS, ANTHERAXANTHIN AND VIOLAXANTHIN, INTO CAPSANTHIN
CC       AND CAPSORUBIN, RESPECTIVELY.
CC   -!- PATHWAY: CAROTENOID BIOSYNTHESIS.
CC   -!- SUBUNIT: MONOMER.
CC   -!- SUBCELLULAR LOCATION: CHLOROPLAST; CHROMOPLAST.
CC   -!- SIMILARITY: BELONGS TO THE LYCOPENE CYCLASE FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; X77289; CAA54495.1; -.
DR   EMBL; X76165; CAA53759.1; -.
DR   EMBL; X78030; CAA54961.1; -.
DR   Mendel; 1; Capan;424;1.
DR   Mendel; 7586; Capan;424;7586.
DR   InterPro; IPR000205; -.
KW   Oxidoreductase; NAD; Carotenoid biosynthesis; Chloroplast;
KW   Transit peptide.
FT   TRANSIT       1      ?       CHLOROPLAST (POTENTIAL).
FT   CHAIN         ?    498       CAPSANTHIN/CAPSORUBIN SYNTHASE.
FT   NP_BIND      84    112       NAD (POTENTIAL).
FT   CONFLICT    316    316       A -> R (IN REF. 3).
FT   CONFLICT    378    380       AEA -> LRP (IN REF. 3).
FT   CONFLICT    406    408       PSD -> LRH (IN REF. 3).
FT   CONFLICT    458    498       ELAVLSLYLFGHASNLARLDIVTKCTVPLVKLLGNLAIESL
FT                                -> RTCCTQFVPFWTCL (IN REF. 3).
SQ   SEQUENCE   498 AA;  56658 MW;  4FB25F676A022A98 CRC64;


  Query Match             65.5%;  Score 38;  DB 1;  Length 498;
  Best Local Similarity   77.8%;  Pred. No. 27;
  Matches    7;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       2 LMPFPGPLL 10
           | ||| |||
Db       5 LKPFPSPLL 13


RESULT   8
X_HPBVO
ID   X_HPBVO        STANDARD;      PRT;   154 AA.
AC   P20977;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   01-NOV-1995 (Rel. 32, Last annotation update)
DE   TRANS-ACTIVATING PROTEIN X.
GN   X.
OS   Hepatitis B virus (subtype adw / strain Okinawa/pODW282).
```

```
RL   Science 287:2185-2195(2000).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=OREGON-R;
RX   MEDLINE=20196011; PubMed=10731137;
RA   Benos P.V., Gatt M.K., Ashburner M., Murphy L., Harris D.,
RA   Barrell B.G., Ferraz C., Vidal S., Brun C., Demailles J., Cadieu E.,
RA   Dreano S., Gloux S., Lelaure V., Mottier S., Galibert F., Borkova D.,
RA   Minana B., Kafatos F.C., Louis C., Siden-Kiamos I., Bolshakov S.,
RA   Papagiannakis G., Spanos L., Cox S., Madueno E., de Pablos B.,
RA   Modolell J., Peter A., Schottler P., Werner M., Mourkioti F.,
RA   Beinert N., Dowe G., Schafer U., Jackle H., Bucheton A.,
RA   Callister D.M., Campbell L.A., Darlamitsou A., Henderson N.S.,
RA   McMillan P.J., Salles C., Tait E.A., Valenti P., Saunder R.D.,
RA   Glover D.M.;
RT   "From sequence to chromosome: the tip of the X chromosome of D.
RT   melanogaster.";
RL   Science 287:2220-2222(2000).
RN   [5]
RP   FUNCTION.
RX   MEDLINE=90059894; PubMed=2583094;
RA   Torres M., Sanchez L.;
RT   "The scute (T4) gene acts as a numerator element of the X: a signal
RT   that determines the state of activity of sex-lethal in Drosophila.";
RL   EMBO J. 8:3079-3086(1989).
CC   -!- FUNCTION: AS-C PROTEINS ARE INVOLVED IN THE DETERMINATION OF THE
CC       NEURONAL PRECURSORS IN THE PERIPHERAL NERVOUS SYSTEM AND THE
CC       CENTRAL NERVOUS SYSTEM. ALSO INVOLVED IN SEX DETERMINATION AND
CC       DOSAGE COMPENSATION.
CC   -!- SUBUNIT: EFFICIENT DNA BINDING REQUIRES DIMERIZATION WITH ANOTHER
CC       BHLH PROTEIN.
CC   -!- TISSUE SPECIFICITY: L(1)SC, SC AND AC STRONGLY LABEL THE
CC       PRESUMPTIVE STOMATOGASTRIC NERVOUS SYSTEM, WHILE ASE IS MORE
CC       PROMINENT IN THE PRESUMPTIVE PROCEPHALIC LOBE.
CC   -!- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF
CC       TRANSCRIPTION FACTORS.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; M17119; AAA28313.1; -.
DR   EMBL; AE003417; AAF45499.1; -.
DR   EMBL; AL024453; CAA19657.1; -.
DR   HSSP; P10085; 1MDY.
DR   TRANSFAC; T00004; -.
DR   FlyBase; FBgn0004170; sc.
DR   InterPro; IPR001092; -.
DR   InterPro; IPR003015; -.
DR   Pfam; PF00010; HLH; 1.
DR   PROSITE; PS00038; HELIX_LOOP_HELIX; 1.
KW   Neurogenesis; Differentiation; Developmental protein.
FT   DNA_BIND     102    112       BASIC DOMAIN.
FT   DOMAIN       113    163       HELIX-LOOP-HELIX MOTIF (BY SIMILARITY).
FT   CONFLICT     161    161       R -> S (IN REF. 1).
FT   CONFLICT     213    213       T -> R (IN REF. 1).
FT   CONFLICT     219    219       L -> V (IN REF. 1).
SQ   SEQUENCE   345 AA;  38155 MW;  DE68E49A8CCF16EB CRC64;

  Query Match             69.0%;  Score 40;  DB 1;  Length 345;
  Best Local Similarity   60.0%;  Pred. No. 8.6;
  Matches    6;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy        1 HLMPFPGPLL 10
            |:|| | ||:
Db       39 HIMPAPSPLI 48
```

```
RESULT   5
ARX_MOUSE
ID   ARX_MOUSE      STANDARD;      PRT;   563 AA.
AC   O35085;
DT   15-JUL-1999 (Rel. 38, Created)
DT   15-JUL-1999 (Rel. 38, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   HOMEOBOX PROTEIN ARX.
GN   ARX.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97398450; PubMed=9256348;
RA   Miura H., Yanazawa M., Kato K., Kitamura K.;
RT   "Expression of a novel aristaless related homeobox gene 'Arx' in the
RT   vertebrate telencephalon, diencephalon and floor plate.";
RL   Mech. Dev. 65:99-109(1997).
CC   -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE PAIRED FAMILY OF HOMEOBOX PROTEINS.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AB006103; BAA28284.1; -.
DR   MGD; MGI:1097716; Arx.
DR   InterPro; IPR000047; -.
DR   InterPro; IPR000104; -.
DR   InterPro; IPR001356; -.
DR   Pfam; PF00046; homeobox; 1.
DR   PRINTS; PR00024; HOMEOBOX.
DR   PRINTS; PR00031; HTHREPRESSR.
DR   PRINTS; PR00308; ANTIFREEZEI.
DR   PROSITE; PS00027; HOMEOBOX_1; 1.
DR   PROSITE; PS50071; HOMEOBOX_2; 1.
KW   Homeobox; DNA-binding; Nuclear protein.
FT   DNA_BIND     330    389       HOMEOBOX.
FT   DOMAIN       100    116       POLY-ALA.
FT   DOMAIN       127    132       POLY-PRO.
FT   DOMAIN       143    154       POLY-ALA.
FT   DOMAIN       183    186       POLY-PRO.
FT   DOMAIN       231    241       POLY-GLU.
FT   DOMAIN       274    285       POLY-ALA.
FT   DOMAIN       434    441       POLY-ALA.
FT   DOMAIN       446    449       POLY-PRO.
SQ   SEQUENCE   563 AA;  58504 MW;  C4CDAA004CA22DFD CRC64;

  Query Match             67.2%;  Score 39;  DB 1;  Length 563;
  Best Local Similarity   85.7%;  Pred. No. 21;
  Matches    6;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        3 MPFPGPL 9
            :||||||
Db      400 LPFPGPL 406


RESULT   6
ENV_BAEVM
ID   ENV_BAEVM      STANDARD;      PRT;   563 AA.
AC   P10269;
DT   01-MAR-1989 (Rel. 10, Created)
DT   01-MAR-1989 (Rel. 10, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   ENV POLYPROTEIN PRECURSOR (COAT POLYPROTEIN) [CONTAINS: OUTER MEMBRANE
DE   PROTEIN GP70; TRANSMEMBRANE PROTEIN P20E].
```

```
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 HLMPFPGPLL 10
            ||||||||||
Db      378 HLMPFPGPLL 387


RESULT   3
GAT1_RAT
ID   GAT1_RAT       STANDARD;      PRT;   413 AA.
AC   P43429;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   ERYTHROID TRANSCRIPTION FACTOR (GATA-1) (ERYF1) (GF-1) (NF-E1).
GN   GATA1.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=WISTAR; TISSUE=Liver;
RX   MEDLINE=95046273; PubMed=7957872;
RA   Matsuda K., Kobune Y., Noda C., Ichihara A.;
RT   "Expression of GATA-binding transcription factors in rat
RT   hepatocytes.";
RL   FEBS Lett. 353:269-272(1994).
CC   -!- FUNCTION: TRANSCRIPTIONAL ACTIVATOR WHICH PROBABLY SERVES AS A
CC       GENERAL SWITCH FACTOR FOR ERYTHROID DEVELOPMENT. IT BINDS TO DNA
CC       SITES WITH THE CONSENSUS SEQUENCE (A/T)GATA(A/G) WITHIN REGULATORY
CC       REGIONS OF GLOBIN GENES AND OF OTHER GENES EXPRESSED IN ERYTHROID
CC       CELLS.
CC   -!- SUBUNIT: THE AMINO-TERMINAL ZINC FINGER INTERACTS WITH FOG-1 (BY
CC       SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: NUCLEAR.
CC   -!- TISSUE SPECIFICITY: ERYTHROCYTE.
CC   -!- DOMAIN: THE TWO FINGERS ARE FUNCTIONALLY DISTINCT AND COOPERATE TO
CC       ACHIEVE SPECIFIC, STABLE DNA BINDING. THE FIRST FINGER IS
CC       NECESSARY ONLY FOR FULL SPECIFICITY AND STABILITY OF BINDING,
CC       WHEREAS THE SECOND ONE IS REQUIRED FOR BINDING (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE GATA-TYPE TRANSCRIPTION FACTOR FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; D13518; BAA02735.1; -.
DR   HSSP; P17678; 1GAU.
DR   InterPro; IPR000679; -.
DR   Pfam; PF00320; GATA; 2.
DR   PRINTS; PR00619; GATAZNFINGER.
DR   PROSITE; PS00344; GATA_ZN_FINGER_1; 2.
DR   PROSITE; PS50114; GATA_ZN_FINGER_2; 2.
KW   Transcription regulation; Activator; DNA-binding; Erythrocyte;
KW   Zinc-finger; Nuclear protein.
FT   ZN_FING     204    228       GATA-TYPE.
FT   ZN_FING     258    282       GATA-TYPE.
FT   BINDING     203    203       FOG-1 (BY SIMILARITY).
FT   BINDING     205    205       FOG-1 (BY SIMILARITY).
FT   BINDING     208    208       FOG-1 (BY SIMILARITY).
FT   BINDING     222    222       FOG-1 (BY SIMILARITY).
SQ   SEQUENCE   413 AA;  42871 MW;  9BE4120FCF8269BB CRC64;


  Query Match             86.2%;  Score 50;  DB 1;  Length 413;
  Best Local Similarity   90.0%;  Pred. No. 0.22;
  Matches    9;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;
```

```
Qy        1 HLMPFPGPLL 10
            ||| ||||||
Db      378 HLMSFPGPLL 387


RESULT   4
AST4_DROME
ID   AST4_DROME     STANDARD;      PRT;   345 AA.
AC   P10084; O76890;
DT   01-MAR-1989 (Rel. 10, Created)
DT   01-OCT-2000 (Rel. 40, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   ACHAETE-SCUTE COMPLEX PROTEIN T4 (SCUTE PROTEIN).
GN   SC OR T4 OR EG:198A6.1 OR CG3827.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CANTON-S;
RX   MEDLINE=87273503; PubMed=3111716;
RA   Villares R., Cabrera C.V.;
RT   "The achaete-scute gene complex of D. melanogaster: conserved domains
RT   in a subset of genes required for neurogenesis and their homology to
RT   myc.";
RL   Cell 50:415-424(1987).
RN   [2]
RP   REVISIONS.
RA   Villares R.;
RL   Submitted (SEP-1988) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BERKELEY;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
```

```
CC       DYSERYTHROPOIETIC ANEMIA AND THROMBOCYTOPENIA (XDAT), A DISORDER
CC       CHARACTERIZED BY ERYTHROCYTES WITH ABNORMAL SIZE AND SHAPE, AND
CC       PAUCITY OF PLATELETS IN PERIPHERAL BLOOD. THE BONE MARROW CONTAINS
CC       ABUNDANT AND ABNORMALLY SMALL MEGAKARYOCYTES.
CC   -!- SIMILARITY: BELONGS TO THE GATA-TYPE TRANSCRIPTION FACTOR FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; M30601; AAA35885.1; -.
DR   EMBL; X17254; CAA35120.1; -.
DR   EMBL; AF196971; AAF06806.1; -.
DR   PIR; A34888; A34888.
DR   PIR; S07121; S07121.
DR   HSSP; P17678; 1GAU.
DR   TRANSFAC; T00227; -.
DR   TRANSFAC; T00306; -.
DR   MIM; 305371; -.
DR   InterPro; IPR000679; -.
DR   Pfam; PF00320; GATA; 2.
DR   PRINTS; PR00619; GATAZNFINGER.
DR   PROSITE; PS00344; GATA_ZN_FINGER_1; 2.
DR   PROSITE; PS50114; GATA_ZN_FINGER_2; 2.
KW   Transcription regulation; Activator; DNA-binding; Erythrocyte;
KW   Zinc-finger; Nuclear protein; Disease mutation.
FT   ZN_FING     204    228       GATA-TYPE.
FT   ZN_FING     258    282       GATA-TYPE.
FT   BINDING     203    203       FOG-1.
FT   BINDING     205    205       FOG-1.
FT   BINDING     208    208       FOG-1.
FT   BINDING     222    222       FOG-1.
FT   VARIANT     205    205       V -> M (IN XDAT; SEVERE IMPAIRMENT OF
FT                                FOG-1 BINDING AND ERYTHROID
FT                                DIFFERENTIATION IN VITRO).
FT                                /FTId=VAR_010115.
FT   MUTAGEN     204    204       C->R: INCREASE OF DISSOCIATION RATE FROM
FT                                BOUND DNA.
SQ   SEQUENCE   413 AA;  42751 MW;  822BD2DE14B908AD CRC64;

  Query Match             100.0%;  Score 58;  DB 1;  Length 413;
  Best Local Similarity   100.0%;  Pred. No. 0.0099;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 HLMPFPGPLL 10
              ||||||||||
Db        378 HLMPFPGPLL 387


RESULT 2
GAT1_MOUSE
ID   GAT1_MOUSE     STANDARD;      PRT;   413 AA.
AC   P17679;
DT   01-AUG-1990 (Rel. 15, Created)
DT   01-AUG-1990 (Rel. 15, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   ERYTHROID TRANSCRIPTION FACTOR (GATA-1) (ERYF1) (GF-1) (NF-E1).
GN   GATA1 OR GF-1.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC   TISSUE=Erythrocyte;
RX   MEDLINE=89262083; PubMed=2725678;
RA   Tsai S.-F., Martin D.I.K., Zon L.I., D'Andrea A.D., Wong G.W.,
RA   Orkin S.H.;
RT   "Cloning of cDNA for the major DNA-binding protein of the erythroid
RT   lineage through expression in mammalian cells.";
RL   Nature 339:446-451(1989).
RN   [2]
RP   SEQUENCE OF 1-73 FROM N.A.
RC   STRAIN=BALB/C;
RA   Todokoro K., Chiba T., Kuramochi S., Ikawa Y.;
RL   Submitted (MAY-1995) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   DETAILED STUDIES OF THE TWO FINGERS.
RX   MEDLINE=91115078; PubMed=2276623;
RA   Martin D.I.K., Orkin S.H.;
RT   "Transcriptional activation and DNA binding by the erythroid factor
RT   GF-1/NF-E1/Eryf 1.";
RL   Genes Dev. 4:1886-1898(1990).
RN   [4]
RP   STRUCTURE BY NMR OF 200-243.
RX   MEDLINE=99229455; PubMed=10212985;
RA   Kowalski K., Czolij R., King G.F., Crossley M., Mackay J.P.;
RT   "The solution structure of the N-terminal zinc finger of GATA-1
RT   reveals a specific binding face for the transcriptional co-factor
RT   FOG.";
RL   J. Biomol. NMR 13:249-262(1999).
CC   -!- FUNCTION: TRANSCRIPTIONAL ACTIVATOR WHICH PROBABLY SERVES AS A
CC       GENERAL SWITCH FACTOR FOR ERYTHROID DEVELOPMENT. IT BINDS TO DNA
CC       SITES WITH THE CONSENSUS SEQUENCE (A/T)GATA(A/G) WITHIN REGULATORY
CC       REGIONS OF GLOBIN GENES AND OF OTHER GENES EXPRESSED IN ERYTHROID
CC       CELLS.
CC   -!- SUBUNIT: THE AMINO-TERMINAL ZINC FINGER INTERACTS WITH FOG-1.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR.
CC   -!- TISSUE SPECIFICITY: ERYTHROCYTE.
CC   -!- DOMAIN: THE TWO FINGERS ARE FUNCTIONALLY DISTINCT AND COOPERATE TO
CC       ACHIEVE SPECIFIC, STABLE DNA BINDING. THE FIRST FINGER IS
CC       NECESSARY ONLY FOR FULL SPECIFICITY AND STABILITY OF BINDING,
CC       WHEREAS THE SECOND ONE IS REQUIRED FOR BINDING.
CC   -!- SIMILARITY: BELONGS TO THE GATA-TYPE TRANSCRIPTION FACTOR FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; X15763; CAA33769.1; -.
DR   EMBL; X57530; CAA40751.1; -.
DR   PIR; S04655; S04655.
DR   PDB; 1GNF; 08-JUN-99.
DR   TRANSFAC; T00305; -.
DR   MGD; MGI:95661; Gata1.
DR   InterPro; IPR000679; -.
DR   Pfam; PF00320; GATA; 2.
DR   PRINTS; PR00619; GATAZNFINGER.
DR   PROSITE; PS00344; GATA_ZN_FINGER_1; 2.
DR   PROSITE; PS50114; GATA_ZN_FINGER_2; 2.
KW   Transcription regulation; Activator; DNA-binding; Erythrocyte;
KW   Zinc-finger; Nuclear protein; 3D-structure.
FT   ZN_FING     204    228       GATA-TYPE.
FT   ZN_FING     258    282       GATA-TYPE.
FT   BINDING     203    203       FOG-1.
FT   BINDING     205    205       FOG-1.
FT   BINDING     208    208       FOG-1.
FT   BINDING     222    222       FOG-1.
FT   MUTAGEN     207    207       C->P: STABILITY OF BINDING REDUCED.
FT   MUTAGEN     230    230       L->F: STABILITY OF BINDING REDUCED.
FT   MUTAGEN     261    261       C->P: ABOLISHES DNA-BINDING.
FT   MUTAGEN     284    284       L->F: BINDS WITH REDUCED AFFINITY.
SQ   SEQUENCE   413 AA;  42674 MW;  BB627A92700D557A CRC64;

  Query Match             100.0%;  Score 58;  DB 1;  Length 413;
  Best Local Similarity   100.0%;  Pred. No. 0.0099;
```

```
                    GenCore version 4.5
          Copyright (c) 1993 - 2000  Compugen Ltd.

OM protein - protein search, using sw model

Run on:          October 23, 2001, 13:28:18 ; Search time 17.76 Seconds
                                              (without alignments)
                                              19.288 Million cell updates/sec

Title:           US-09-625-963-3
Perfect score:   58
Sequence:        1 HLMPFPGPLL 10

Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched:        93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters:      93435

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       SwissProt_39:*

     Pred. No. is the number of results predicted by chance to have a
     score greater than or equal to the score of the result being printed,
     and is derived by analysis of the total score distribution.

                               SUMMARIES
                   %
Result           Query
  No.   Score  Match Length DB  ID                    Description
 -------------------------------------------------------------------------
    1      58  100.0    413  1  GAT1_HUMAN            P15976 homo sapien
    2      58  100.0    413  1  GAT1_MOUSE            P17679 mus musculu
    3      50   86.2    413  1  GAT1_RAT              P43429 rattus norv
    4      40   69.0    345  1  AST4_DROME            P10084 drosophila
    5      39   67.2    563  1  ARX_MOUSE             O35085 mus musculu
    6      39   67.2    563  1  ENV_BAEVM             P10269 baboon endo
    7      38   65.5    498  1  CCS_CAPAN             Q42435 capsicum an
    8      37   63.8    154  1  X_HPBVO               P20977 hepatitis b
    9      37   63.8    778  1  TAST_HUMAN            Q12815 homo sapien
   10      36   62.1    389  1  NDPP_MOUSE            Q03173 mus musculu
   11      36   62.1    486  1  EGR4_HUMAN            Q05215 homo sapien
   12      36   62.1    738  1  VNUC_EBOSB            Q9qp77 ebola virus
   13      36   62.1    739  1  VNUC_EBOG4            Q9qce9 ebola virus
   14      36   62.1    739  1  VNUC_EBOZ5            O72142 ebola virus
   15      36   62.1    739  1  VNUC_EBOZM            P18272 ebola virus
   16      35   60.3     83  1  UR2_PLAFE             P21857 platichthys
   17      35   60.3    224  1  CASB_BOVIN            P02666 bos taurus
   18      35   60.3    224  1  CASB_BUBBU            Q9tsi0 bubalus bub
   19      35   60.3    462  1  ERR1_MOUSE            O08580 mus musculu
   20      35   60.3    504  1  A37C_DROME            P18487 drosophila
   21      35   60.3    518  1  CP3R_ONCMY            O42563 oncorhynchu
   22      35   60.3    519  1  ERR1_HUMAN            P11474 homo sapien
   23      35   60.3    568  1  VL1_HPV18             P06794 human papil
   24      35   60.3    575  1  ENV_SMRVH             P21412 squirrel mo
   25      35   60.3    690  1  NCPR_PHAAU            P37116 phaseolus a
   26    34.5   59.5    516  1  ER11_ARATH            O65404 arabidopsis
   27      34   58.6     13  1  CRBL_ICASP            P17237 icaria sp.
   28      34   58.6    183  1  AR21_CAEEL            Q9xwv3 caenorhabdi
   29      34   58.6    188  1  LCAT_PIG              P30930 sus scrofa
   30      34   58.6    246  1  Y419_ARCFU            O29828 archaeoglob
   31      34   58.6    270  1  TSNR_STRLU            P52393 streptomyce
   32      34   58.6    290  1  COBD_ARCFU            O28933 archaeoglob
   33      34   58.6    311  1  Y4FM_RHISN            P55451 rhizobium s
   34      34   58.6    324  1  YM02_MYCTU            Q10391 mycobacteri
   35      34   58.6    337  1  FXL1_MOUSE            Q64731 mus musculu
   36      34   58.6    377  1  CYCI_HUMAN            Q14094 homo sapien
   37      34   58.6    377  1  CYCI_MOUSE            Q9z2v9 mus musculu
   38      34   58.6    384  1  SX18_HUMAN            P35713 homo sapien
   39      34   58.6    450  1  GNUT_PSEAE            Q9zij1 pseudomonas
   40      34   58.6    506  1  ER11_BRANA            O65727 brassica na
   41      34   58.6    517  1  ER12_ARATH            O65402 arabidopsis
   42      34   58.6    672  1  BGAL_BACST            P19668 bacillus st
   43      34   58.6   2214  1  SORL_HUMAN            Q92673 h sortilin-
   44      33   56.9    127  1  KRCL_CHICK            P25692 gallus gall
   45      33   56.9    154  1  KRSC_CHICK            P04459 gallus gall

                          ALIGNMENTS

RESULT   1
GAT1_HUMAN
ID   GAT1_HUMAN     STANDARD;      PRT;   413 AA.
AC   P15976;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   ERYTHROID TRANSCRIPTION FACTOR (GATA-1) (ERYF1) (GF-1) (NF-E1).
GN   GATA1 OR GF1 OR ERYF1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC   TISSUE=Erythrocyte;
RX   MEDLINE=90138889; PubMed=2300555;
RA   Zon L.I., Tsai S.-F., Burgess S., Matsudaira P., Bruns G.A.P.,
RA   Orkin S.H.;
RT   "The major human erythroid DNA-binding protein (GF-1): primary
RT   sequence and localization of the gene to the X chromosome.";
RL   Proc. Natl. Acad. Sci. U.S.A. 87:668-672(1990).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Bone marrow;
RX   MEDLINE=90114418; PubMed=2104960;
RA   Trainor C.D., Evans T., Felsenfeld G., Boguski M.S.;
RT   "Structure and evolution of a human erythroid transcription factor.";
RL   Nature 343:92-96(1990).
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Blechschmidt K., Nyakatura G., Strom T.M., Drescher B., Menzel U.,
RA   Meindl A., Rosenthal A.;
RL   Submitted (OCT-1999) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   FOG-1 BINDING, VARIANT XDAT MET-205, AND MUTAGENESIS OF CYS-204.
RC   TISSUE=Peripheral blood;
RX   MEDLINE=20164324; PubMed=10700180;
RA   Nichols K.E., Crispino J.D., Poncz M., White J.G., Orkin S.H.,
RA   Maris J.M., Weiss M.J.;
RT   "Familial dyserythropoietic anaemia and thrombocytopenia due to an
RT   inherited mutation in GATA1.";
RL   Nat. Genet. 24:266-270(2000).
CC   -!- FUNCTION: TRANSCRIPTIONAL ACTIVATOR WHICH PROBABLY SERVES AS A
CC       GENERAL SWITCH FACTOR FOR ERYTHROID DEVELOPMENT. IT BINDS TO DNA
CC       SITES WITH THE CONSENSUS SEQUENCE (A/T)GATA(A/G) WITHIN REGULATORY
CC       REGIONS OF GLOBIN GENES AND OF OTHER GENES EXPRESSED IN ERYTHROID
CC       CELLS.
CC   -!- SUBUNIT: THE AMINO-TERMINAL ZINC FINGER INTERACTS WITH FOG-1.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR.
CC   -!- TISSUE SPECIFICITY: ERYTHROCYTE.
CC   -!- DOMAIN: THE TWO FINGERS ARE FUNCTIONALLY DISTINCT AND COOPERATE TO
CC       ACHIEVE SPECIFIC, STABLE DNA BINDING. THE FIRST FINGER IS
CC       NECESSARY ONLY FOR FULL SPECIFICITY AND STABILITY OF BINDING,
CC       WHEREAS THE SECOND ONE IS REQUIRED FOR BINDING (BY SIMILARITY).
CC   -!- DISEASE: DEFECTS IN GATA1 ARE THE CAUSE OF X-LINKED
```

```
RESULT  15
JS0603
gene X protein - hepatitis B virus (subtype adw, strain Okinawa/pODW282)
C;Species: hepatitis B virus, HBV
C;Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 14-Nov-1997
C;Accession: JS0603
R;Okamoto, H.; Tsuda, F.; Sakugawa, H.; Sastrosoewignjo, R.I.; Imai, M.; Miyakawa, Y.; M
J. Gen. Virol. 69, 2575-2583, 1988
A;Title: Typing hepatitis B virus by homology in nucleotide sequence: comparison of surf
A;Reference number: JS0253; MUID:89010694
A;Accession: JS0603
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-154 <OKA>
A;Cross-references: GB:D00330; NID:g221498
C;Genetics:
A;Gene: X
C;Superfamily: hepatitis B virus gene X protein


  Query Match             63.8%;  Score 37;  DB 2;  Length 154;
  Best Local Similarity   100.0%;  Pred. No. 23;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       4 PFPGPL 9
           ||||||
Db      29 PFPGPL 34


Search completed: October 23, 2001, 13:22:07
Job time: 120 sec
```

```
Qy       3 MPFPGP 8
           ||||||
Db     186 MPFPGP 191


RESULT  11
F83094
cell division protein FtsW PA4413 [imported] - Pseudomonas aeruginosa (strain PAO1)
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C;Accession: F83094
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrener, P.; Hickey, M.J.; Br
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A;Reference number: A82950; MUID:20437337
A;Accession: F83094
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-399 <STO>
A;Cross-references: GB:AE004856; GB:AE004091; NID:g9950633; PIDN:AAG07801.1; GSPDB:GN001
A;Experimental source: strain PAO1
C;Genetics:
A;Gene: ftsW; PA4413
C;Superfamily: rod shape-determining protein


  Query Match             65.5%;  Score 38;  DB 2;  Length 399;
  Best Local Similarity   77.8%;  Pred. No. 42;
  Matches    7;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       2 LMPFPGPLL 10
           | ||| |||
Db       5 LRPFPSPLL 13

RESULT  12
S51511
capsanthin/capsorubin synthase (EC 5.5.1.-) - pepper
C;Species: Capsicum annuum (pepper)
C;Date: 19-Mar-1997 #sequence_revision 25-Apr-1997 #text_change 28-Jul-2000
C;Accession: S51511
R;Houlne, G.; Schantz, M.L.; Meyer, B.; Pozueta-Romero, J.; Schantz, R.
Curr. Genet. 26, 524-527, 1994
A;Title: A chromoplast-specific protein in Capsicum annuum: characterization and express
A;Reference number: S51511; MUID:95179811
A;Accession: S51511
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-471 <HOU>
A;Cross-references: GB:X78030; NID:g840728; PIDN:CAA54961.1; PID:g840729
C;Superfamily: tomato lycopene beta-cyclase
C;Keywords: intramolecular lyase; isomerase


  Query Match             65.5%;  Score 38;  DB 2;  Length 471;
  Best Local Similarity   77.8%;  Pred. No. 50;
  Matches    7;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       2 LMPFPGPLL 10
           | ||| |||
Db       5 LKPFPSPLL 13

RESULT  13
S71511
capsanthin/capsorubin synthase (EC 5.5.1.-) - pepper
N;Alternate names: chromoplast-specific protein
N;Contains: capsorubin synthase; capssanthin synthase
C;Species: Capsicum annuum (pepper)
C;Date: 09-Dec-1997 #sequence_revision 09-Dec-1997 #text_change 28-Jul-2000
C;Accession: S71511; JC2141
R;Bouvier, F.; Hugueney, P.; d'Harlingue, A.; Kuntz, M.; Camara, B.
Plant J. 6, 45-54, 1994
A;Title: Xanthophyll biosynthesis in chromoplasts: isolation and molecular cloning of
A;Reference number: S71511; MUID:95004653
A;Accession: S71511
A;Molecule type: mRNA
A;Residues: 1-498 <BOU>
A;Cross-references: EMBL:X76165; NID:g522119; PIDN:CAA53759.1; PID:g522120
A;Experimental source: cv. Lamuyo; developmental stage ripening fruits
R;Deruere, J.; Bouvier, F.; Steppuhn, J.; Klein, A.; Camara, B.; Kuntz, M.
Biochem. Biophys. Res. Commun. 199, 1144-1150, 1994
A;Title: Structure and expression of two plant genes encoding chromoplast-specific pr
A;Reference number: JC2140; MUID:94197697
A;Accession: JC2141
A;Molecule type: DNA
A;Residues: 1-498 <DER>
A;Cross-references: EMBL:X77289; NID:g468747; PIDN:CAA54495.1; PID:g468748
A;Experimental source: cv. Yolo Wonder
R;Deruere, J.; Bouvier, F.; Steppuhn, J.; Klein, A.; Camara, B.; Kuntz, M.
Biochem. Biophys. Res. Commun. 201, 486, 1994
A;Reference number: A55775
A;Contents: annotation; erratum
A;Note: the legends of the nucleotide sequence figures were reversed in publication (
C;Genetics:
A;Gene: CCS
C;Complex: monomer
C;Function:
A;Description: catalyses the conversion of the ubiquitous 5,6-epoxicartenoids, anther
A;Pathway: carotenoid biosynthesis
C;Superfamily: tomato lycopene beta-cyclase
C;Keywords: chromoplast; intramolecular lyase; isomerase


  Query Match             65.5%;  Score 38;  DB 2;  Length 498;
  Best Local Similarity   77.8%;  Pred. No. 53;
  Matches    7;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       2 LMPFPGPLL 10
           | ||| |||
Db       5 LKPFPSPLL 13

RESULT  14
S74743
hypothetical protein sll0922 - Synechocystis sp. (strain PCC 6803)
C;Species: Synechocystis sp.
A;Variety: PCC 6803
C;Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 20-Jun-2000
C;Accession: S74743
R;Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima,
o, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yas
DNA Res. 3, 109-136, 1996
A;Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocys
s.
A;Reference number: S74322; MUID:97061201
A;Accession: S74743
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-116 <KAN>
A;Cross-references: EMBL:D90901; GB:AB001339; NID:g1651897; PIDN:BAA16894.1; PID:g165
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
C;Genetics:
A;Start codon: GTG
C;Superfamily: Synechocystis hypothetical protein sll0922


  Query Match             63.8%;  Score 37;  DB 2;  Length 116;
  Best Local Similarity   66.7%;  Pred. No. 17;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 LMPFPGPLL 10
           ::| |||||
Db       1 MIPVPGPLL 9
```

```
A;Note: Host Papio sp. (baboon)
C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 16-Jul-1999
C;Accession: JT0262
R;Kato, S.; Matsuo, K.; Nishimura, N.; Takahashi, N.; Takano, T.
Jpn. J. Genet. 62, 127-137, 1987
A;Title: The entire nucleotide sequence of baboon endogenous virus DNA: a chimeric genom
A;Reference number: JT0260
A;Accession: JT0262
A;Molecule type: DNA
A;Residues: 1-563 <KAT>
A;Cross-references: GB:M16550; NID:g509586; PIDN:AAA87333.1; PID:g332599
C;Genetics:
A;Gene: env
C;Superfamily: type C retrovirus env polyprotein
C;Keywords: coat protein; glycoprotein; polyprotein; transmembrane protein
F;1-376/Product: coat protein gp70 #status predicted <CGP>
F;377-563/Product: coat protein p20E #status predicted <CPP>
F;437-449/Region: immunosuppressive peptide #status predicted
F;113,219,229,264,282,292,306,312,321,339,469/Binding site: carbohydrate (Asn) (covalent


  Query Match             67.2%;  Score 39;  DB 1;  Length 563;
  Best Local Similarity   70.0%;  Pred. No. 41;
  Matches    7;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 HLMPFPGPLL 10
             :|:|| ||||
Db     507 YLLPFLGPLL 516


RESULT   7
E69281
GMP synthase (guaA-1) homolog - Archaeoglobus fulgidus
C;Species: Archaeoglobus fulgidus
C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 22-Oct-1999
C;Accession: E69281
R;Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson
.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.
 Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
Nature 390, 364-370, 1997
A;Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.
Smith, H.O.; Woese, C.R.; Venter, J.C.
A;Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaec
A;Reference number: A69250; MUID:98049343
A;Accession: E69281
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-303 <KLE>
A;Cross-references: GB:AE001088; GB:AE000782; NID:g2689411; PIDN:AAB90977.1; PID:g265038


  Query Match             65.5%;  Score 38;  DB 2;  Length 303;
  Best Local Similarity   100.0%;  Pred. No. 32;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3 MPFPGP 8
             ||||||
Db     184 MPFPGP 189

RESULT   8
A84334
GMP synthase subunit B [imported] - Halobacterium sp. NRC-1
C;Species: Halobacterium sp. NRC-1
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C;Accession: A84334
R;Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S
; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablo
 Jung, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A;Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li
A;Title: Genome sequence of Halobacterium species NRC-1.
A;Reference number: A84160; MUID:20504483
```

```
A;Accession: A84334
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-305 <STO>
A;Cross-references: GB:AE004437; NID:g10581275; PIDN:AAG20037.1; GSPDB:GN00138
C;Genetics:
A;Gene: guaAb


  Query Match             65.5%;  Score 38;  DB 2;  Length 305;
  Best Local Similarity   100.0%;  Pred. No. 32;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3 MPFPGP 8
             ||||||
Db     186 MPFPGP 191


RESULT   9
C75125
gmp synthase, cter domain PAB2417 - Pyrococcus abyssi (strain Orsay)
C;Species: Pyrococcus abyssi
C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Aug-1999
C;Accession: C75125
R;anonymous, Genoscope
submitted to the EMBL Data Library, July 1999
A;Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome s
A;Reference number: A75001
A;Accession: C75125
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-308 <KAW>
A;Cross-references: GB:AJ248285; GB:AL096836; NID:g5458067; PIDN:CAB49716.1; PID:e151
A;Experimental source: strain Orsay
C;Genetics:
A;Gene: guaA-C; PAB2417


  Query Match             65.5%;  Score 38;  DB 2;  Length 308;
  Best Local Similarity   100.0%;  Pred. No. 32;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3 MPFPGP 8
             ||||||
Db     186 MPFPGP 191


RESULT  10
E71006
probable GMP synthase - Pyrococcus horikoshii
C;Species: Pyrococcus horikoshii
C;Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 05-Nov-1999
C;Accession: E71006
R;Kawarabayasi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Se
 M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Ogu
DNA Res. 5, 55-76, 1998
A;Title: Complete sequence and gene organization of the genome of a hyper-thermophili
A;Reference number: A71000; MUID:98344137
A;Accession: E71006
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-308 <KAW>
A;Cross-references: GB:AP000006; NID:g3236133; PIDN:BAA30453.1; PID:d1031396; PID:g32
A;Experimental source: strain OT3
A;Note: this accession replaces an interim accession for a sequence replaced by GenBa
C;Genetics:
A;Gene: PH1347


  Query Match             65.5%;  Score 38;  DB 2;  Length 308;
  Best Local Similarity   100.0%;  Pred. No. 32;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
C;Date: 28-Feb-1990 #sequence_revision 28-Feb-1990 #text_change 26-Aug-1999
C;Accession: S04655; S55394
R;Tsai, S.F.; Martin, D.I.K.; Zon, L.I.; D'Andrea, A.D.; Wong, G.G.; Orkin, S.H.
Nature 339, 446-451, 1989
A;Title: Cloning of cDNA for the major DNA-binding protein of the erythroid lineage thr
A;Reference number: S04655; MUID:89262083
A;Accession: S04655
A;Status: not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-413 <TSA>
A;Cross-references: GB:X15763; NID:g51063; PIDN:CAA33769.1; PID:g51064
R;Todokoro, K.; Chiba, T.; Kuramochi, S.; Ikawa, Y.
submitted to the EMBL Data Library, February 1991
A;Description: Nucleotide sequences of the 5'-flanking region of the mouse GATA-1 gene
A;Reference number: S55394
A;Accession: S55394
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-73 <TOD>
A;Cross-references: EMBL:X57530; NID:g854279; PIDN:CAA40751.1; PID:g854280
C;Superfamily: transcription factor GATA-1; GATA-type zinc finger homology
C;Keywords: DNA binding; nucleus; transcription factor; transcription regulation; zinc
F;201-254/Domain: GATA-type zinc finger homology <GZF1>
F;204-228/Region: zinc finger GATA motif
F;255-308/Domain: GATA-type zinc finger homology <GZF2>
F;258-282/Region: zinc finger GATA motif


  Query Match             100.0%;  Score 58;  DB 2;  Length 413;
  Best Local Similarity   100.0%;  Pred. No. 0.017;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 HLMPFPGPLL 10
                ||||||||||
Db        378 HLMPFPGPLL 387


RESULT   3
S48756
transcription factor GATA-1 - rat
N;Alternate names: finger protein GF-1; transcription factor Eryf1
C;Species: Rattus norvegicus (Norway rat)
C;Date: 27-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 20-Jun-2000
C;Accession: S48756; S35891
R;Matsuda, K.; Kobune, Y.; Noda, C.; Ichihara, A.
FEBS Lett. 353, 269-272, 1994
A;Title: Expression of GATA-binding transcription factors in rat hepatocytes.
A;Reference number: S48756; MUID:95046273
A;Accession: S48756
A;Molecule type: mRNA
A;Residues: 1-413 <MAT>
A;Cross-references: EMBL:D13518; NID:g644895; PIDN:BAA02735.1; PID:g644896
R;Noda, C.
submitted to the EMBL Data Library, October 1992
A;Reference number: S35891
A;Accession: S35891
A;Molecule type: mRNA
A;Residues: 1-279,'KC',282-413 <NOD>
A;Cross-references: EMBL:D13518
C;Superfamily: transcription factor GATA-1; GATA-type zinc finger homology
C;Keywords: DNA binding; nucleus; transcription factor; transcription regulation; zinc
F;201-254/Domain: GATA-type zinc finger homology <GZF1>
F;255-308/Domain: GATA-type zinc finger homology <GZF2>


  Query Match             86.2%;  Score 50;  DB 2;  Length 413;
  Best Local Similarity   90.0%;  Pred. No. 0.39;
  Matches    9;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 HLMPFPGPLL 10
                ||| ||||||
Db        378 HLMSFPGPLL 387
```

```
RESULT   4
F83608
conserved hypothetical protein PA0292 [imported] - Pseudomonas aeruginosa (strain PAO
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C;Accession: F83608
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrener, P.; Hickey, M.J.;
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; L
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pa
A;Reference number: A82950; MUID:20437337
A;Accession: F83608
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-368 <STO>
A;Cross-references: GB:AE004467; GB:AE004091; NID:g9946133; PIDN:AAG03681.1; GSPDB:GN
A;Experimental source: strain PAO1
C;Genetics:
A;Gene: PA0292


  Query Match             72.4%;  Score 42;  DB 2;  Length 368;
  Best Local Similarity   77.8%;  Pred. No. 8.1;
  Matches    7;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 HLMPFPGPL 9
                | || ||||
Db        267 HKMPIPGPL 275


RESULT   5
B43731
achaete-scute complex protein T4 - fruit fly (Drosophila melanogaster)
C;Species: Drosophila melanogaster
C;Date: 03-Mar-1993 #sequence_revision 12-Mar-1993 #text_change 17-Mar-1999
C;Accession: B43731; S35425
R;Villares, R.; Cabrera, C.V.
Cell 50, 415-424, 1987
A;Title: The achaete-scute gene complex of Drosophila melanogaster: conserved domains
A;Reference number: A43731; MUID:87273503
A;Accession: B43731
A;Molecule type: DNA
A;Residues: 1-345 <VIL1>
A;Cross-references: GB:M17119
R;Villares, R.
submitted to the EMBL Data Library, November 1990
A;Reference number: S35425
A;Accession: S35425
A;Molecule type: DNA
A;Residues: 1-255,'C',257-345 <VIL2>
A;Cross-references: EMBL:M17119; NID:g156745; PID:g156748
C;Genetics:
A;Gene: FlyBase:sc
A;Cross-references: FlyBase:FBgn0004170


  Query Match             69.0%;  Score 40;  DB 2;  Length 345;
  Best Local Similarity   60.0%;  Pred. No. 17;
  Matches    6;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 HLMPFPGPLL 10
                |:|| | ||:
Db         39 HIMPAPSPLI 48


RESULT   6
VCMVM7
env polyprotein - baboon endogenous virus (strain M7)
N;Alternate names: coat polyprotein
N;Contains: coat protein gp70; coat protein p20E
C;Species: baboon endogenous virus
```

```
                    GenCore version 4.5
          Copyright (c) 1993 - 2000  Compugen Ltd.

OM protein - protein search, using sw model

Run on:         October 23, 2001, 13:22:06 ; Search time 27.18 Seconds
                                              (without alignments)
                                              28.026 Million cell updates/sec

Title:          US-09-625-963-3
Perfect score:  58
Sequence:       1 HLMPFPGPLL 10

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched:       219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters:      219241

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       PIR_68:*
                1:  pir1:*
                2:  pir2:*
                3:  pir3:*
                4:  pir4:*

     Pred. No. is the number of results predicted by chance to have a
     score greater than or equal to the score of the result being printed,
     and is derived by analysis of the total score distribution.

                              SUMMARIES
                 %
Result         Query
  No.   Score  Match Length DB  ID                          Description
-------------------------------------------------------------------------
    1      58  100.0    413  1  A34888                      transcription fact
    2      58  100.0    413  2  S04655                      transcription fact
    3      50   86.2    413  2  S48756                      transcription fact
    4      42   72.4    368  2  F83608                      conserved hypothet
    5      40   69.0    345  2  B43731                      achaete-scute comp
    6      39   67.2    563  1  VCMVM7                      env polyprotein -
    7      38   65.5    303  2  E69281                      GMP synthase (guaA
    8      38   65.5    305  2  A84334                      GMP synthase subun
    9      38   65.5    308  2  C75125                      gmp synthase, cter
   10      38   65.5    308  2  E71006                      probable GMP synth
   11      38   65.5    399  2  F83094                      cell division prot
   12      38   65.5    471  2  S51511                      capsanthin/capsoru
   13      38   65.5    498  2  S71511                      capsanthin/capsoru
   14      37   63.8    116  2  S74743                      hypothetical prote
   15      37   63.8    154  2  JS0603                      gene X protein - h
   16      37   63.8    778  2  I38487                      tastin - human
   17      37   63.8    833  2  G75621                      arylesterase/monox
   18      36   62.1    343  2  T28908                      hypothetical prote
   19      36   62.1    362  2  T40377                      hypothetical prote
   20      36   62.1    389  2  S27200                      proline-rich prote
   21      36   62.1    486  2  A41537                      DNA-binding protei
   22      36   62.1    739  1  VHIWEB                      nucleocapsid prote
   23      36   62.1    954  2  T24748                      hypothetical prote
   24      36   62.1   1000  2  T21522                      hypothetical prote
   25      35   60.3      6  2  A61049                      halo-toxin - Pseud
   26      35   60.3     29  2  S77569                      plantaricin SA6 -
   27      35   60.3     83  2  S10706                      urotensin II precu
   28      35   60.3    100  2  T43116                      hypothetical prote
   29      35   60.3    102  2  G71025                      hypothetical prote
   30      35   60.3    203  2  E75343                      hypothetical prote
   31      35   60.3    209  2  A59068                      beta-casein varian
   32      35   60.3    224  1  KBBOA2                      beta-casein precur
   33      35   60.3    294  2  H84023                      phosphate ABC tran
   34      35   60.3    334  2  T36284                      probable araC-fami
   35      35   60.3    338  2  T35540                      probable kinase -
   36      35   60.3    347  2  H75253                      hypothetical prote
   37      35   60.3    521  2  A29345                      steroid hormone re
   38      35   60.3    568  1  P1WL18                      L1 protein - human
   39      35   60.3    575  1  VCLJHD                      env polyprotein pr
   40      35   60.3    686  2  B75061                      alpha chain of for
   41      35   60.3    690  2  A47298                      NADPH--ferrihemopr
   42      35   60.3    692  2  S37159                      NADPH--ferrihemopr
   43      35   60.3    706  2  H82764                      O-antigen acetylas
   44      35   60.3    720  2  T02457                      hypothetical prote
   45      35   60.3    774  2  B84031                      ATP-dependent prot

                              ALIGNMENTS

RESULT   1
A34888
transcription factor GATA-1 - human
N;Alternate names: finger protein GF-1; transcription factor Eryf1; transcription fac
C;Species: Homo sapiens (man)
C;Date: 20-Jul-1990 #sequence_revision 12-Jul-1996 #text_change 16-Jul-1999
C;Accession: A34888; S07121
R;Zon, L.I.; Tsai, S.F.; Burgess, S.; Matsudaira, P.; Bruns, G.A.P.; Orkin, S.H.
Proc. Natl. Acad. Sci. U.S.A. 87, 668-672, 1990
A;Title: The major human erythroid DNA-binding protein (GF-1): primary sequence and l
A;Reference number: A34888; MUID:90138889
A;Accession: A34888
A;Molecule type: mRNA
A;Residues: 1-413 <ZON>
A;Cross-references: GB:M30601; NID:g183071; PIDN:AAA35885.1; PID:g183072
R;Trainor, C.D.; Evans, T.; Felsenfeld, G.; Boguski, M.S.
Nature 343, 92-96, 1990
A;Title: Structure and evolution of a human erythroid transcription factor.
A;Reference number: S07121; MUID:90114418
A;Accession: S07121
A;Molecule type: mRNA
A;Residues: 1-413 <TRA>
A;Cross-references: EMBL:X17254; NID:g31242; PIDN:CAA35120.1; PID:g31243
C;Comment: This transcriptional activator is named for the core, GATA, of the nucleot
C;Comment: GATA-1 appears to be a major regulator of both globin and non-globin eryth
C;Comment: GATA-1 is expressed in erythroid cells, megakaryocytes, and mast cells of
C;Genetics:
A;Gene: GDB:GATA1; GF1
A;Cross-references: GDB:125373; OMIM:305371
A;Map position: Xp11.23-Xp11.23
C;Superfamily: transcription factor GATA-1; GATA-type zinc finger homology
C;Keywords: DNA binding; nucleus; transcription factor; transcription regulation; zin
F;201-254/Domain: GATA-type zinc finger homology <GZF1>
F;204-228/Region: zinc finger GATA motif
F;255-308/Domain: GATA-type zinc finger homology <GZF2>
F;258-282/Region: zinc finger GATA motif


  Query Match             100.0%;  Score 58;  DB 1;  Length 413;
  Best Local Similarity   100.0%;  Pred. No. 0.017;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 HLMPFPGPLL 10
            ||||||||||
Db      378 HLMPFPGPLL 387


RESULT   2
S04655
transcription factor GATA-1 - mouse
N;Alternate names: finger protein GF-1; protein EF-1; protein EF-gamma-a; protein Ery
C;Species: Mus musculus (house mouse)
```

```
  Matches     5;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 CMTWNQMNL 9
           |: || | |
Db      29 CLPWNSMIL 37


RESULT  15
US-08-427-640-2
; Sequence 2, Application US/08427640
; Patent No. 5658788
;  GENERAL INFORMATION:
;    APPLICANT:  Berg et al.
;    TITLE OF INVENTION:  Compounds and Methods for Treatment of Thromboembolic Disor
;    NUMBER OF SEQUENCES:  28
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Eli Lilly and Company
;      STREET:  Lilly Corporate Center
;      CITY:  Indianapolis
;      STATE:  IN.
;      COUNTRY:  U.S.A.
;      ZIP:  46285
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.50 inch, 1.0 Mb storage
;      COMPUTER:  Macintosh
;      OPERATING SYSTEM:  Macintosh
;      SOFTWARE:  Microsoft Word
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/427,640
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/689,410
;      FILING DATE:  22 APRIL 1991
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  355 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-427-640-2


  Query Match             58.9%;  Score 33;  DB 1;  Length 355;
  Best Local Similarity   55.6%;  Pred. No. 1.9e+02;
  Matches     5;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 CMTWNQMNL 9
           |: || | |
Db      29 CLPWNSMIL 37


Search completed: October 23, 2001, 13:21:32
Job time: 85 sec
```

```
;      ADDRESSEE:  OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
;      ADDRESSEE:  P.C.
;      STREET:  1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
;      CITY:  ARLINGTON
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22202
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/811,949
;      FILING DATE:  05-MAR-1997
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  OBLON, NORMAN F.
;      REGISTRATION NUMBER:  24,618
;      REFERENCE/DOCKET NUMBER:  18-966-0
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-413-3000
;      TELEFAX:  703-413-2220
;  INFORMATION FOR SEQ ID NO:  61:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  354 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-811-949-61


  Query Match             58.9%;  Score 33;  DB 2;  Length 354;
  Best Local Similarity   55.6%;  Pred. No. 1.9e+02;
  Matches    5;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 CMTWNQMNL 9
           |: || | |
Db      29 CLPWNSMIL 37


RESULT  13
US-08-137-116-1
; Sequence 1, Application US/08137116
; Patent No. 5500411
;  GENERAL INFORMATION:
;    APPLICANT:  Martin, Ulrich
;    APPLICANT:  Koenig, Reinhard
;    TITLE OF INVENTION:  METHOD FOR TREATING THROMBOEMBOLIC
;    TITLE OF INVENTION:  CONDITIONS BY INHIBITING REOCCLUSION
;    TITLE OF INVENTION:  VIA THE USE OF MULTIPLE BOLUS
;    TITLE OF INVENTION:  ADMINISTRATION OF THROMBOLYTICALLY
;    TITLE OF INVENTION:  ACTIVE PROTEINS
;    NUMBER OF SEQUENCES:  2
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Felfe & Lynch
;      STREET:  805 Third Avenue
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10022
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 5.25 inch, 360 kb storage
;      COMPUTER:  IBM PS/2
;      OPERATING SYSTEM:  PC-DOS
;      SOFTWARE:  Wordperfect
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/137,116
;      FILING DATE:  June 30, 1994
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/EP92/00851
;      FILING DATE:  15 April 1992
```

```
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  DE 41 12 398
;      FILING DATE:  16 April 1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  DE 41 23 845
;      FILING DATE:  18 July 1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Hanson, No. 5500411man D.
;      REGISTRATION NUMBER:  30,946
;      REFERENCE/DOCKET NUMBER:  BOER 1026
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 688-9200
;      TELEFAX:  (212) 838-3884
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  355 amino acids
;      TYPE:  amino acids
;      TOPOLOGY:  linear
US-08-137-116-1

  Query Match             58.9%;  Score 33;  DB 1;  Length 355;
  Best Local Similarity   55.6%;  Pred. No. 1.9e+02;
  Matches    5;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 CMTWNQMNL 9
           |: || | |
Db      29 CLPWNSMIL 37

RESULT  14
US-08-217-618-1
; Sequence 1, Application US/08217618
; Patent No. 5510330
;  GENERAL INFORMATION:
;    APPLICANT:  Martin, Ulrich
;    APPLICANT:  Fischer, Stephan
;    TITLE OF INVENTION:  COMBINATIONS OF THROMBOLYTICALLY ACTIVE
;    TITLE OF INVENTION:  PROTEINS AND NON-HEPARIN ANTICOAGULANTS, AND USES THEREOF
;    NUMBER OF SEQUENCES:  1
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Felfe & Lynch
;      STREET:  805 Third Avenue
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10022
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 5.25 inch, 360 kb storage
;      COMPUTER:  IBM PS/2
;      OPERATING SYSTEM:  PC-DOS
;      SOFTWARE:  Wordperfect
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/217,618
;      FILING DATE:  25-MARCH-1994
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Hanson, No. 5510330man D.
;      REGISTRATION NUMBER:  30,946
;      REFERENCE/DOCKET NUMBER:  BOER 1032
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 688-9200
;      TELEFAX:  (212) 838-3884
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  355 amino acids
;      TYPE:  amino acids
;      TOPOLOGY:  linear
US-08-217-618-1

  Query Match             58.9%;  Score 33;  DB 1;  Length 355;
  Best Local Similarity   55.6%;  Pred. No. 1.9e+02;
```

```
;      LENGTH:  837 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  NO
US-09-012-710-12


  Query Match             60.7%;  Score 34;  DB 3;  Length 837;
  Best Local Similarity   50.0%;  Pred. No. 3.1e+02;
  Matches    4;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        1 CMTWNQMN 8
            |::|:| |
Db      496 CVSWSQFN 503


RESULT  10
US-08-664-596B-18
; Sequence 18, Application US/08664596B
; Patent No. 5807703
;  GENERAL INFORMATION:
;    APPLICANT:  Jacobs, Kenneth
;    APPLICANT:  McCoy, John
;    APPLICANT:  LaVallie, Edward
;    APPLICANT:  Racie, Lisa
;    APPLICANT:  Merberg, David
;    APPLICANT:  Treacy, Maurice
;    APPLICANT:  Evans, Cheryl
;    APPLICANT:  Spaulding, Vikki
;    APPLICANT:  Bowman, Michael
;    TITLE OF INVENTION:  SECRETED PROTEINS AND POLYNUCLEOTIDES
;    TITLE OF INVENTION:  ENCODING THEM
;    NUMBER OF SEQUENCES:  37
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genetics Institute, Inc.
;      STREET:  87 CambridgePark Drive
;      CITY:  Cambridge
;      STATE:  Massachusetts
;      COUNTRY:  U.S.A.
;      ZIP:  02140
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/664,596B
;      FILING DATE:
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Brown, Scott A.
;      REGISTRATION NUMBER:  32,724
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 498-8224
;      TELEFAX:  (617) 876-5851
;  INFORMATION FOR SEQ ID NO:  18:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  131 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-664-596B-18


  Query Match             58.9%;  Score 33;  DB 1;  Length 131;
  Best Local Similarity   75.0%;  Pred. No. 66;
  Matches    6;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy        1 CMTWNQMN 8
            ||| || |
```

```
Db       17 CMTTNQTN 24

RESULT  11
US-08-811-949-1
; Sequence 1, Application US/08811949
; Patent No. 5840533
;  GENERAL INFORMATION:
;    APPLICANT:  NIWA, MINEO
;    APPLICANT:  SAITO, YOSHIMASA
;    APPLICANT:  SASAKI, HITOSHI
;    APPLICANT:  HAYASHI, MASAKO
;    APPLICANT:  NOTANI, JOUJI
;    APPLICANT:  KOBAYASHI, MASAKAZU
;    TITLE OF INVENTION:  TISSUE PLASMINOGEN ACTIVATOR
;    NUMBER OF SEQUENCES:  67
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
;      ADDRESSEE:  P.C.
;      STREET:  1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
;      CITY:  ARLINGTON
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22202
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/811,949
;      FILING DATE:  05-MAR-1997
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  OBLON, NORMAN F.
;      REGISTRATION NUMBER:  24,618
;      REFERENCE/DOCKET NUMBER:  18-966-0
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-413-3000
;      TELEFAX:  703-413-2220
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  347 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-811-949-1


  Query Match             58.9%;  Score 33;  DB 2;  Length 347;
  Best Local Similarity   55.6%;  Pred. No. 1.8e+02;
  Matches    5;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy        1 CMTWNQMNL 9
            |: || | |
Db       28 CLPWNSMIL 36


RESULT  12
US-08-811-949-61
; Sequence 61, Application US/08811949
; Patent No. 5840533
;  GENERAL INFORMATION:
;    APPLICANT:  NIWA, MINEO
;    APPLICANT:  SAITO, YOSHIMASA
;    APPLICANT:  SASAKI, HITOSHI
;    APPLICANT:  HAYASHI, MASAKO
;    APPLICANT:  NOTANI, JOUJI
;    APPLICANT:  KOBAYASHI, MASAKAZU
;    TITLE OF INVENTION:  TISSUE PLASMINOGEN ACTIVATOR
;    NUMBER OF SEQUENCES:  67
;    CORRESPONDENCE ADDRESS:
```

```
; Patent No. 5726288
;  GENERAL INFORMATION:
;    APPLICANT:  Call, Katherine M.
;    APPLICANT:  Glaser, Thomas M.
;    APPLICANT:  Ito, Caryn Y.
;    APPLICANT:  Buckler, Alan J.
;    APPLICANT:  Pelletier, Jerry
;    APPLICANT:  Haber, Daniel A.
;    APPLICANT:  Rose, Elise A.
;    APPLICANT:  Housman, David E.
;    APPLICANT:  Bruening, Wendy
;    APPLICANT:  Darveau, Andre
;    TITLE OF INVENTION:  Localization and Characterization of the
;    TITLE OF INVENTION:  Wilms' Tumor Gene
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Hamilton, Brook, Smith & Reynolds, P.C.
;      STREET:  Two Militia Drive
;      CITY:  Lexington
;      STATE:  Massachusetts
;      COUNTRY:  U.S.
;      ZIP:  02173
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/102,942A
;      FILING DATE:  02-AUG-1993
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Granahan, Patricia
;      REGISTRATION NUMBER:  32,227
;      REFERENCE/DOCKET NUMBER:  MIT-5194A2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-861-6240
;      TELEFAX:  617-861-9540
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  154 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  not relevant
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-102-942A-5

  Query Match             73.2%;  Score 41;  DB 1;  Length 154;
  Best Local Similarity   100.0%;  Pred. No. 3.6;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 CMTWNQ 6
          ||||||
Db    149 CMTWNQ 154


RESULT   8
US-08-906-791-2
; Sequence 2, Application US/08906791
; Patent No. 6096876
;  GENERAL INFORMATION:
;    APPLICANT:  St-Arnaud, Rene
;    APPLICANT:  Glorieux, Francis H.
;    TITLE OF INVENTION:  1-alpha-Hydroxylase Materials and
;    TITLE OF INVENTION:  Methods
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  6300 Sears Towers, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  U.S.A.
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/906,791
;      FILING DATE:  06-AUG-1997
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  O'Toole, Edward M.
;      REGISTRATION NUMBER:  22477
;      REFERENCE/DOCKET NUMBER:  29105/34166
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (312) 474-6300
;      TELEFAX:  (312) 474-0448
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  501 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-906-791-2

  Query Match             60.7%;  Score 34;  DB 3;  Length 501;
  Best Local Similarity   71.4%;  Pred. No. 1.8e+02;
  Matches    5;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 CMTWNQM 7
          |  ||||
Db    247 CRDWNQM 253


RESULT   9
US-09-012-710-12
; Sequence 12, Application US/09012710
; Patent No. 6087478
;  GENERAL INFORMATION:
;    APPLICANT:  Vinkemeier, Uwe
;    APPLICANT:  Moarefi, Ismail
;    APPLICANT:  Darnell, Jr., James E.
;    APPLICANT:  Kuriyan, John
;    TITLE OF INVENTION:  A CRYSTAL OF THE N-TERMINAL DOMAIN OF A
;    TITLE OF INVENTION:  STAT PROTEIN AND METHODS OF USE THEREOF
;    NUMBER OF SEQUENCES:  13
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Klauber & Jackson
;      STREET:  411 Hackensack Avenue, 4th Floor
;      CITY:  Hackensack
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/012,710
;      FILING DATE:
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jackson Esq., David A.
;      REGISTRATION NUMBER:  26,742
;      REFERENCE/DOCKET NUMBER:  600-1-194
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201-487-5800
;      TELEFAX:  201-343-1684
;      TELEX:  133521
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
```

```
;      NAME:  Bak, Mary E.
;      REGISTRATION NUMBER:  31,215
;      REFERENCE/DOCKET NUMBER:  WST48PCT
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  215-540-9200
;      TELEFAX:  215-540-5818
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  429 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
PCT-US95-05523-4


  Query Match             100.0%;  Score 56;  DB 5;  Length 429;
  Best Local Similarity   100.0%;  Pred. No. 0.033;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CMTWNQMNL 9
           |||||||||
Db     235 CMTWNQMNL 243


RESULT   5
US-08-102-942A-4
; Sequence 4, Application US/08102942A
; Patent No. 5726288
;  GENERAL INFORMATION:
;    APPLICANT:  Call, Katherine M.
;    APPLICANT:  Glaser, Thomas M.
;    APPLICANT:  Ito, Caryn Y.
;    APPLICANT:  Buckler, Alan J.
;    APPLICANT:  Pelletier, Jerry
;    APPLICANT:  Haber, Daniel A.
;    APPLICANT:  Rose, Elise A.
;    APPLICANT:  Housman, David E.
;    APPLICANT:  Bruening, Wendy
;    APPLICANT:  Darveau, Andre
;    TITLE OF INVENTION:  Localization and Characterization of the
;    TITLE OF INVENTION:  Wilms' Tumor Gene
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Hamilton, Brook, Smith & Reynolds, P.C.
;      STREET:  Two Militia Drive
;      CITY:  Lexington
;      STATE:  Massachusetts
;      COUNTRY:  U.S.
;      ZIP:  02173
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/102,942A
;      FILING DATE:  02-AUG-1993
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Granahan, Patricia
;      REGISTRATION NUMBER:  32,227
;      REFERENCE/DOCKET NUMBER:  MIT-5194A2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-861-6240
;      TELEFAX:  617-861-9540
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  449 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-102-942A-4
```

```
  Query Match             100.0%;  Score 56;  DB 1;  Length 449;
  Best Local Similarity   100.0%;  Pred. No. 0.034;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CMTWNQMNL 9
           |||||||||
Db     235 CMTWNQMNL 243


RESULT   6
US-08-102-942A-6
; Sequence 6, Application US/08102942A
; Patent No. 5726288
;  GENERAL INFORMATION:
;    APPLICANT:  Call, Katherine M.
;    APPLICANT:  Glaser, Thomas M.
;    APPLICANT:  Ito, Caryn Y.
;    APPLICANT:  Buckler, Alan J.
;    APPLICANT:  Pelletier, Jerry
;    APPLICANT:  Haber, Daniel A.
;    APPLICANT:  Rose, Elise A.
;    APPLICANT:  Housman, David E.
;    APPLICANT:  Bruening, Wendy
;    APPLICANT:  Darveau, Andre
;    TITLE OF INVENTION:  Localization and Characterization of the
;    TITLE OF INVENTION:  Wilms' Tumor Gene
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Hamilton, Brook, Smith & Reynolds, P.C.
;      STREET:  Two Militia Drive
;      CITY:  Lexington
;      STATE:  Massachusetts
;      COUNTRY:  U.S.
;      ZIP:  02173
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/102,942A
;      FILING DATE:  02-AUG-1993
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Granahan, Patricia
;      REGISTRATION NUMBER:  32,227
;      REFERENCE/DOCKET NUMBER:  MIT-5194A2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-861-6240
;      TELEFAX:  617-861-9540
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  449 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  not relevant
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-102-942A-6


  Query Match             100.0%;  Score 56;  DB 1;  Length 449;
  Best Local Similarity   100.0%;  Pred. No. 0.034;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CMTWNQMNL 9
           |||||||||
Db     235 CMTWNQMNL 243


RESULT   7
US-08-102-942A-5
; Sequence 5, Application US/08102942A
```

```
  Query Match             100.0%;  Score 56;  DB 1;  Length 345;
  Best Local Similarity   100.0%;  Pred. No. 0.026;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CMTWNQMNL 9
           |||||||||
Db     151 CMTWNQMNL 159


RESULT   2
US-08-234-783-4
; Sequence 4, Application US/08234783
; Patent No. 5622835
;  GENERAL INFORMATION:
;    APPLICANT:  Herlyn, Meenhard
;    APPLICANT:  Morris, Jennifer
;    APPLICANT:  Rauscher III, Frank J.
;    APPLICANT:  Rodeck, Ulrich
;    TITLE OF INVENTION:  WT1 Monoclonal Antibodies and Methods of
;    TITLE OF INVENTION:  Use Therefor
;    NUMBER OF SEQUENCES:  7
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Howson and Howson
;      STREET:  Spring House Corporate Cntr, PO Box 457
;      CITY:  Spring House
;      STATE:  Pennsylvania
;      COUNTRY:  USA
;      ZIP:  19477
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/234,783
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Bak, Mary E.
;      REGISTRATION NUMBER:  31,215
;      REFERENCE/DOCKET NUMBER:  WST48USA
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  215-540-9200
;      TELEFAX:  215-540-5818
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  429 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-234-783-4


  Query Match             100.0%;  Score 56;  DB 1;  Length 429;
  Best Local Similarity   100.0%;  Pred. No. 0.033;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CMTWNQMNL 9
           |||||||||
Db     235 CMTWNQMNL 243


RESULT   3
US-08-456-907-4
; Sequence 4, Application US/08456907
; Patent No. 5633142
;  GENERAL INFORMATION:
;    APPLICANT:  Herlyn, Meenhard
;    APPLICANT:  Morris, Jennifer
;    APPLICANT:  Rauscher III, Frank J.
;    APPLICANT:  Rodeck, Ulrich
;    TITLE OF INVENTION:  WT1 Monoclonal Antibodies and Methods of
;    TITLE OF INVENTION:  Use Therefor
;    NUMBER OF SEQUENCES:  7
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Howson and Howson
;      STREET:  Spring House Corporate Cntr, PO Box 457
;      CITY:  Spring House
;      STATE:  Pennsylvania
;      COUNTRY:  USA
;      ZIP:  19477
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/456,907
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/234,783
;      FILING DATE:  28-APR-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Bak, Mary E.
;      REGISTRATION NUMBER:  31,215
;      REFERENCE/DOCKET NUMBER:  WST48AUSA
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  215-540-9200
;      TELEFAX:  215-540-5818
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  429 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-456-907-4


  Query Match             100.0%;  Score 56;  DB 1;  Length 429;
  Best Local Similarity   100.0%;  Pred. No. 0.033;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CMTWNQMNL 9
           |||||||||
Db     235 CMTWNQMNL 243


RESULT   4
PCT-US95-05523-4
; Sequence 4, Application PC/TUS9505523
;  GENERAL INFORMATION:
;    APPLICANT:  The Wistar Institute of Anatomy and Biology
;    TITLE OF INVENTION:  WT1 Monoclonal Antibodies and
;    TITLE OF INVENTION:  Methods of Use Therefor
;    NUMBER OF SEQUENCES:  7
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Howson and Howson
;      STREET:  Spring House Corporate Cntr, PO Box 457
;      CITY:  Spring House
;      STATE:  Pennsylvania
;      COUNTRY:  USA
;      ZIP:  19477
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/05523
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/234,783
;      FILING DATE:  28-APR-1994
;    ATTORNEY/AGENT INFORMATION:
```

```
                     GenCore version 4.5
          Copyright (c) 1993 - 2000  Compugen Ltd.


OM protein - protein search, using sw model

Run on:         October 23, 2001, 13:21:32 ; Search time 23.71 Seconds
                                             (without alignments)
                                             7.816 Million cell updates/sec

Title:          US-09-625-963-2
Perfect score:  56
Sequence:       1 CMTWNQMNL 9

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched:       197339 seqs, 20590346 residues

Total number of hits satisfying chosen parameters:      197339

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       Issued_Patents_AA:*
                1:  /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
                2:  /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
                3:  /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
                4:  /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
                5:  /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
                6:  /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                             SUMMARIES
                  %
Result          Query
  No.   Score   Match Length DB  ID                        Description
-------------------------------------------------------------------------
     1     56   100.0    345  1  US-08-102-942A-2           Sequence 2, Appli
     2     56   100.0    429  1  US-08-234-783-4            Sequence 4, Appli
     3     56   100.0    429  1  US-08-456-907-4            Sequence 4, Appli
     4     56   100.0    429  5  PCT-US95-05523-4           Sequence 4, Appli
     5     56   100.0    449  1  US-08-102-942A-4           Sequence 4, Appli
     6     56   100.0    449  1  US-08-102-942A-6           Sequence 6, Appli
     7     41    73.2    154  1  US-08-102-942A-5           Sequence 5, Appli
     8     34    60.7    501  3  US-08-906-791-2            Sequence 2, Appli
     9     34    60.7    837  3  US-09-012-710-12           Sequence 12, Appl
    10     33    58.9    131  1  US-08-664-596B-18          Sequence 18, Appl
    11     33    58.9    347  2  US-08-811-949-1            Sequence 1, Appli
    12     33    58.9    354  2  US-08-811-949-61           Sequence 61, Appl
    13     33    58.9    355  1  US-08-137-116-1            Sequence 1, Appli
    14     33    58.9    355  1  US-08-217-618-1            Sequence 1, Appli
    15     33    58.9    355  1  US-08-427-640-2            Sequence 2, Appli
    16     33    58.9    355  1  US-08-427-640-6            Sequence 6, Appli
    17     33    58.9    355  1  US-08-217-617A-1           Sequence 1, Appli
    18     33    58.9    355  1  US-08-217-616-1            Sequence 1, Appli
    19     33    58.9    355  2  US-08-811-949-45           Sequence 45, Appl
    20     33    58.9    355  2  US-08-811-949-47           Sequence 47, Appl
    21     33    58.9    355  2  US-08-811-949-53           Sequence 53, Appl
    22     33    58.9    355  2  US-08-811-949-59           Sequence 59, Appl
    23     33    58.9    355  4  US-08-794-528-1            Sequence 1, Appli
    24     33    58.9    355  6  5223256-1                 Patent No. 5223256
    25     33    58.9    389  2  US-08-811-949-65           Sequence 65, Appl
    26     33    58.9    389  2  US-08-811-949-67           Sequence 67, Appl
    27     33    58.9    437  2  US-08-811-949-49           Sequence 49, Appl
```

```
    28     33    58.9    437  2  US-08-811-949-51           Sequence 51, Appl
    29     33    58.9    437  2  US-08-811-949-55           Sequence 55, Appl
    30     33    58.9    437  2  US-08-811-949-57           Sequence 57, Appl
    31     33    58.9    472  2  US-08-811-949-63           Sequence 63, Appl
    32     33    58.9    527  1  US-07-609-510B-16          Sequence 16, Appl
    33     33    58.9    527  2  US-08-811-949-39           Sequence 39, Appl
    34     33    58.9    527  5  PCT-US91-01025A-2          Sequence 2, Appli
    35     33    58.9    527  6  5185259-8                 Patent No. 5185259
    36     33    58.9    527  6  5520913-1                 Patent No. 5520913
    37     33    58.9    546  6  5200340-6                 Patent No. 5200340
    38     33    58.9    562  2  US-08-811-949-43           Sequence 43, Appl
    39     33    58.9    562  2  US-08-560-098A-50          Sequence 50, Appl
    40     33    58.9    562  2  US-08-883-795A-38          Sequence 38, Appl
    41     33    58.9    562  6  5185259-3                 Patent No. 5185259
    42     33    58.9    562  6  5200340-2                 Patent No. 5200340
    43     33    58.9    562  6  5244676-5                 Patent No. 5244676
    44     33    58.9    562  6  5344773-2                 Patent No. 5344773
    45     31    55.4    135  3  US-08-946-914-16           Sequence 16, Appl


                             ALIGNMENTS


RESULT   1
US-08-102-942A-2
; Sequence 2, Application US/08102942A
; Patent No. 5726288
;  GENERAL INFORMATION:
;    APPLICANT:  Call, Katherine M.
;    APPLICANT:  Glaser, Thomas M.
;    APPLICANT:  Ito, Caryn Y.
;    APPLICANT:  Buckler, Alan J.
;    APPLICANT:  Pelletier, Jerry
;    APPLICANT:  Haber, Daniel A.
;    APPLICANT:  Rose, Elise A.
;    APPLICANT:  Housman, David E.
;    APPLICANT:  Bruening, Wendy
;    APPLICANT:  Darveau, Andre
;    TITLE OF INVENTION:  Localization and Characterization of the
;    TITLE OF INVENTION:  Wilms' Tumor Gene
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Hamilton, Brook, Smith & Reynolds, P.C.
;      STREET:  Two Militia Drive
;      CITY:  Lexington
;      STATE:  Massachusetts
;      COUNTRY:  U.S.
;      ZIP:  02173
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/102,942A
;      FILING DATE:  02-AUG-1993
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Granahan, Patricia
;      REGISTRATION NUMBER:  32,227
;      REFERENCE/DOCKET NUMBER:  MIT-5194A2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-861-6240
;      TELEFAX:  617-861-9540
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  345 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;  MOLECULE TYPE:  protein
US-08-102-942A-2
```

```
RESULT  14
US-09-929-315-4
; Sequence 4, Application US/09929315
; GENERAL INFORMATION:
;  APPLICANT: Call, Katherine M.
;  APPLICANT:  Glaser, Thomas M.
;  APPLICANT:  Ito, Caryn Y.
;  APPLICANT:  Buckler, Alan J.
;  APPLICANT:  Pelletier, Jerry
;  APPLICANT:  Haber, Daniel A.
;  APPLICANT:  Rose, Elise A.
;  APPLICANT:  Housman, David E.
;  APPLICANT:  Bruening, Wendy
;  APPLICANT:  Darveau, Andre
;  TITLE OF INVENTION: Localization and Characterization of the
;  TITLE OF INVENTION:  Wilms' Tumor Gene
;  FILE REFERENCE: 0050.1312-013
;  CURRENT APPLICATION NUMBER: US/09/929,315
;  CURRENT FILING DATE:  2001-08-14
;  PRIOR APPLICATION NUMBER: US 09/037,179
;  PRIOR FILING DATE: 1998-03-09
;  PRIOR APPLICATION NUMBER: US 08/102,942
;  PRIOR FILING DATE: 1993-08-02
;  PRIOR APPLICATION NUMBER: US 07/614,161
;  PRIOR FILING DATE: 1990-11-13
;  PRIOR APPLICATION NUMBER: US 07/435,780
;  PRIOR FILING DATE: 1989-11-13
;  PRIOR APPLICATION NUMBER: US 07/795,323
;  PRIOR FILING DATE: 1991-11-20
;  NUMBER OF SEQ ID NOS: 21
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4
;   LENGTH: 449
;   TYPE: PRT
;   ORGANISM: Unknown
;   FEATURE:
;   OTHER INFORMATION: Murine
US-09-929-315-4


  Query Match             100.0%;  Score 56;  DB 5;  Length 449;
  Best Local Similarity   100.0%;  Pred. No. 0.023;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 CMTWNQMNL 9
          |||||||||
Db    235 CMTWNQMNL 243


RESULT  15
US-09-929-315-6
; Sequence 6, Application US/09929315
; GENERAL INFORMATION:
;  APPLICANT: Call, Katherine M.
;  APPLICANT:  Glaser, Thomas M.
;  APPLICANT:  Ito, Caryn Y.
;  APPLICANT:  Buckler, Alan J.
;  APPLICANT:  Pelletier, Jerry
;  APPLICANT:  Haber, Daniel A.
;  APPLICANT:  Rose, Elise A.
;  APPLICANT:  Housman, David E.
;  APPLICANT:  Bruening, Wendy
;  APPLICANT:  Darveau, Andre
;  TITLE OF INVENTION: Localization and Characterization of the
;  TITLE OF INVENTION:  Wilms' Tumor Gene
;  FILE REFERENCE: 0050.1312-013
;  CURRENT APPLICATION NUMBER: US/09/929,315
;  CURRENT FILING DATE:  2001-08-14
;  PRIOR APPLICATION NUMBER: US 09/037,179
;  PRIOR FILING DATE: 1998-03-09
;  PRIOR APPLICATION NUMBER: US 08/102,942
;  PRIOR FILING DATE: 1993-08-02
;  PRIOR APPLICATION NUMBER: US 07/614,161
;  PRIOR FILING DATE: 1990-11-13
;  PRIOR APPLICATION NUMBER: US 07/435,780
;  PRIOR FILING DATE: 1989-11-13
;  PRIOR APPLICATION NUMBER: US 07/795,323
;  PRIOR FILING DATE: 1991-11-20
;  NUMBER OF SEQ ID NOS: 21
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6
;   LENGTH: 449
;   TYPE: PRT
;   ORGANISM: Homo sapien
US-09-929-315-6


  Query Match             100.0%;  Score 56;  DB 5;  Length 449;
  Best Local Similarity   100.0%;  Pred. No. 0.023;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 CMTWNQMNL 9
          |||||||||
Db    235 CMTWNQMNL 243


Search completed: October 23, 2001, 13:22:51
Job time: 164 sec
```

```
RESULT  10
US-09-938-864-333
; Sequence 333, Application US/09938864
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Smithgall, Molly
;  APPLICANT:  Moulton, Gus
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Sleath, Paul R.
;  APPLICANT:  Mossman, Sally
;  APPLICANT:  Evans, Lawrence
;  APPLICANT:  Spies, A. Gregory
;  APPLICANT:  Boydston, Jeremy
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1
;  TITLE OF INVENTION:  SPECIFIC IMMUNOTHERAPY
;  FILE REFERENCE: 210121.465C5
;  CURRENT APPLICATION NUMBER: US/09/938,864
;  CURRENT FILING DATE:  2001-08-24
;  NUMBER OF SEQ ID NOS: 413
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 333
;   LENGTH: 410
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-938-864-333


  Query Match             100.0%;  Score 56;  DB 5;  Length 410;
  Best Local Similarity   100.0%;  Pred. No. 0.021;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 CMTWNQMNL 9
          |||||||||
Db    396 CMTWNQMNL 404

RESULT  11
US-09-938-864-393
; Sequence 393, Application US/09938864
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Smithgall, Molly
;  APPLICANT:  Moulton, Gus
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Sleath, Paul R.
;  APPLICANT:  Mossman, Sally
;  APPLICANT:  Evans, Lawrence
;  APPLICANT:  Spies, A. Gregory
;  APPLICANT:  Boydston, Jeremy
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1
;  TITLE OF INVENTION:  SPECIFIC IMMUNOTHERAPY
;  FILE REFERENCE: 210121.465C5
;  CURRENT APPLICATION NUMBER: US/09/938,864
;  CURRENT FILING DATE:  2001-08-24
;  NUMBER OF SEQ ID NOS: 413
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 393
;   LENGTH: 420
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-938-864-393


  Query Match             100.0%;  Score 56;  DB 5;  Length 420;
  Best Local Similarity   100.0%;  Pred. No. 0.021;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 CMTWNQMNL 9
          |||||||||
Db    374 CMTWNQMNL 382
```

```
RESULT  12
US-09-938-864-408
; Sequence 408, Application US/09938864
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Smithgall, Molly
;  APPLICANT:  Moulton, Gus
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Sleath, Paul R.
;  APPLICANT:  Mossman, Sally
;  APPLICANT:  Evans, Lawrence
;  APPLICANT:  Spies, A. Gregory
;  APPLICANT:  Boydston, Jeremy
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1
;  TITLE OF INVENTION:  SPECIFIC IMMUNOTHERAPY
;  FILE REFERENCE: 210121.465C5
;  CURRENT APPLICATION NUMBER: US/09/938,864
;  CURRENT FILING DATE:  2001-08-24
;  NUMBER OF SEQ ID NOS: 413
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 408
;   LENGTH: 429
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-938-864-408


  Query Match             100.0%;  Score 56;  DB 5;  Length 429;
  Best Local Similarity   100.0%;  Pred. No. 0.022;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 CMTWNQMNL 9
          |||||||||
Db    235 CMTWNQMNL 243

RESULT  13
US-09-538-092-960
; Sequence 960, Application US/09538092
; GENERAL INFORMATION:
;  APPLICANT: Giot, Loic
;  APPLICANT:  Mansfield, Traci A.
;  TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
;  FILE REFERENCE: 15966-542
;  CURRENT APPLICATION NUMBER: US/09/538,092
;  CURRENT FILING DATE:  2000-03-29
;  PRIOR APPLICATION NUMBER: 60/127,352
;  PRIOR FILING DATE: 1999-04-01
;  PRIOR APPLICATION NUMBER: 60/178,965
;  PRIOR FILING DATE: 2000-02-01
;  NUMBER OF SEQ ID NOS: 1387
;  SOFTWARE: CuraPatSeqFormatter Version 0.9
; SEQ ID NO 960
;   LENGTH: 449
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (0)...(0)
;   OTHER INFORMATION: Polypeptide Accession Number P19544
US-09-538-092-960


  Query Match             100.0%;  Score 56;  DB 5;  Length 449;
  Best Local Similarity   100.0%;  Pred. No. 0.023;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 CMTWNQMNL 9
          |||||||||
Db    235 CMTWNQMNL 243
```

```
;  APPLICANT:  Evans, Lawrence
;  APPLICANT:  Spies, A. Gregory
;  APPLICANT:  Boydston, Jeremy
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1
;  TITLE OF INVENTION:  SPECIFIC IMMUNOTHERAPY
;  FILE REFERENCE: 210121.465C5
;  CURRENT APPLICATION NUMBER: US/09/938,864
;  CURRENT FILING DATE:  2001-08-24
;  NUMBER OF SEQ ID NOS: 413
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 391
;   LENGTH: 344
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-938-864-391


  Query Match             100.0%;  Score 56;  DB 5;  Length 344;
  Best Local Similarity   100.0%;  Pred. No. 0.018;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 CMTWNQMNL 9
          |||||||||
Db    298 CMTWNQMNL 306

RESULT   7
US-09-929-315-2
; Sequence 2, Application US/09929315
; GENERAL INFORMATION:
;  APPLICANT: Call, Katherine M.
;  APPLICANT:  Glaser, Thomas M.
;  APPLICANT:  Ito, Caryn Y.
;  APPLICANT:  Buckler, Alan J.
;  APPLICANT:  Pelletier, Jerry
;  APPLICANT:  Haber, Daniel A.
;  APPLICANT:  Rose, Elise A.
;  APPLICANT:  Housman, David E.
;  APPLICANT:  Bruening, Wendy
;  APPLICANT:  Darveau, Andre
;  TITLE OF INVENTION: Localization and Characterization of the
;  TITLE OF INVENTION:  Wilms' Tumor Gene
;  FILE REFERENCE: 0050.1312-013
;  CURRENT APPLICATION NUMBER: US/09/929,315
;  CURRENT FILING DATE:  2001-08-14
;  PRIOR APPLICATION NUMBER: US 09/037,179
;  PRIOR FILING DATE: 1998-03-09
;  PRIOR APPLICATION NUMBER: US 08/102,942
;  PRIOR FILING DATE: 1993-08-02
;  PRIOR APPLICATION NUMBER: US 07/614,161
;  PRIOR FILING DATE: 1990-11-13
;  PRIOR APPLICATION NUMBER: US 07/435,780
;  PRIOR FILING DATE: 1989-11-13
;  PRIOR APPLICATION NUMBER: US 07/795,323
;  PRIOR FILING DATE: 1991-11-20
;  NUMBER OF SEQ ID NOS: 21
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
;   LENGTH: 345
;   TYPE: PRT
;   ORGANISM: Homo sapien
US-09-929-315-2

  Query Match             100.0%;  Score 56;  DB 5;  Length 345;
  Best Local Similarity   100.0%;  Pred. No. 0.018;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 CMTWNQMNL 9
          |||||||||
Db    151 CMTWNQMNL 159
```

```
RESULT   8
US-09-938-864-394
; Sequence 394, Application US/09938864
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Smithgall, Molly
;  APPLICANT:  Moulton, Gus
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Sleath, Paul R.
;  APPLICANT:  Mossman, Sally
;  APPLICANT:  Evans, Lawrence
;  APPLICANT:  Spies, A. Gregory
;  APPLICANT:  Boydston, Jeremy
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1
;  TITLE OF INVENTION:  SPECIFIC IMMUNOTHERAPY
;  FILE REFERENCE: 210121.465C5
;  CURRENT APPLICATION NUMBER: US/09/938,864
;  CURRENT FILING DATE:  2001-08-24
;  NUMBER OF SEQ ID NOS: 413
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 394
;   LENGTH: 362
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-938-864-394


  Query Match             100.0%;  Score 56;  DB 5;  Length 362;
  Best Local Similarity   100.0%;  Pred. No. 0.018;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 CMTWNQMNL 9
          |||||||||
Db    168 CMTWNQMNL 176

RESULT   9
US-09-938-864-346
; Sequence 346, Application US/09938864
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Smithgall, Molly
;  APPLICANT:  Moulton, Gus
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Sleath, Paul R.
;  APPLICANT:  Mossman, Sally
;  APPLICANT:  Evans, Lawrence
;  APPLICANT:  Spies, A. Gregory
;  APPLICANT:  Boydston, Jeremy
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1
;  TITLE OF INVENTION:  SPECIFIC IMMUNOTHERAPY
;  FILE REFERENCE: 210121.465C5
;  CURRENT APPLICATION NUMBER: US/09/938,864
;  CURRENT FILING DATE:  2001-08-24
;  NUMBER OF SEQ ID NOS: 413
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 346
;   LENGTH: 369
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-938-864-346


  Query Match             100.0%;  Score 56;  DB 5;  Length 369;
  Best Local Similarity   100.0%;  Pred. No. 0.019;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 CMTWNQMNL 9
          |||||||||
Db    175 CMTWNQMNL 183
```

```
;  APPLICANT:  Boydston, Jeremy
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1
;  TITLE OF INVENTION:  SPECIFIC IMMUNOTHERAPY
;  FILE REFERENCE: 210121.465C5
;  CURRENT APPLICATION NUMBER: US/09/938,864
;  CURRENT FILING DATE:  2001-08-24
;  NUMBER OF SEQ ID NOS: 413
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 258
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-938-864-258

  Query Match             100.0%;  Score 56;  DB 5;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.9e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CMTWNQMNL 9
           |||||||||
Db       1 CMTWNQMNL 9

RESULT   3
US-09-938-864-344
; Sequence 344, Application US/09938864
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Smithgall, Molly
;  APPLICANT:  Moulton, Gus
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Sleath, Paul R.
;  APPLICANT:  Mossman, Sally
;  APPLICANT:  Evans, Lawrence
;  APPLICANT:  Spies, A. Gregory
;  APPLICANT:  Boydston, Jeremy
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1
;  TITLE OF INVENTION:  SPECIFIC IMMUNOTHERAPY
;  FILE REFERENCE: 210121.465C5
;  CURRENT APPLICATION NUMBER: US/09/938,864
;  CURRENT FILING DATE:  2001-08-24
;  NUMBER OF SEQ ID NOS: 413
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 344
;   LENGTH: 133
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-938-864-344

  Query Match             100.0%;  Score 56;  DB 5;  Length 133;
  Best Local Similarity   100.0%;  Pred. No. 0.0072;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CMTWNQMNL 9
           |||||||||
Db      31 CMTWNQMNL 39

RESULT   4
US-09-938-864-395
; Sequence 395, Application US/09938864
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Smithgall, Molly
;  APPLICANT:  Moulton, Gus
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Sleath, Paul R.
;  APPLICANT:  Mossman, Sally
;  APPLICANT:  Evans, Lawrence
```

```
;  APPLICANT:  Spies, A. Gregory
;  APPLICANT:  Boydston, Jeremy
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1
;  TITLE OF INVENTION:  SPECIFIC IMMUNOTHERAPY
;  FILE REFERENCE: 210121.465C5
;  CURRENT APPLICATION NUMBER: US/09/938,864
;  CURRENT FILING DATE:  2001-08-24
;  NUMBER OF SEQ ID NOS: 413
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 395
;   LENGTH: 214
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-938-864-395

  Query Match             100.0%;  Score 56;  DB 5;  Length 214;
  Best Local Similarity   100.0%;  Pred. No. 0.011;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CMTWNQMNL 9
           |||||||||
Db     168 CMTWNQMNL 176

RESULT   5
US-09-938-864-335
; Sequence 335, Application US/09938864
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Smithgall, Molly
;  APPLICANT:  Moulton, Gus
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Sleath, Paul R.
;  APPLICANT:  Mossman, Sally
;  APPLICANT:  Evans, Lawrence
;  APPLICANT:  Spies, A. Gregory
;  APPLICANT:  Boydston, Jeremy
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1
;  TITLE OF INVENTION:  SPECIFIC IMMUNOTHERAPY
;  FILE REFERENCE: 210121.465C5
;  CURRENT APPLICATION NUMBER: US/09/938,864
;  CURRENT FILING DATE:  2001-08-24
;  NUMBER OF SEQ ID NOS: 413
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 335
;   LENGTH: 256
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-938-864-335

  Query Match             100.0%;  Score 56;  DB 5;  Length 256;
  Best Local Similarity   100.0%;  Pred. No. 0.013;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CMTWNQMNL 9
           |||||||||
Db     242 CMTWNQMNL 250

RESULT   6
US-09-938-864-391
; Sequence 391, Application US/09938864
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Smithgall, Molly
;  APPLICANT:  Moulton, Gus
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Sleath, Paul R.
;  APPLICANT:  Mossman, Sally
```

```
                    GenCore version 4.5
          Copyright (c) 1993 - 2000  Compugen Ltd.


OM protein - protein search, using sw model

Run on:         October 23, 2001, 13:22:51 ; Search time 39.01 Seconds
                                             (without alignments)
                                             11.888 Million cell updates/sec

Title:          US-09-625-963-2
Perfect score:  56
Sequence:       1 CMTWNQMNL 9

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched:       339663 seqs, 51527020 residues

Total number of hits satisfying chosen parameters:      339663

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       Pending_Patents_AA_New:*
                1:  /cgn2_6/ptodata/2/paa/PCT_NEW_COMB.pep:*
                2:  /cgn2_6/ptodata/2/paa/US06_NEW_COMB.pep:*
                3:  /cgn2_6/ptodata/2/paa/US07_NEW_COMB.pep:*
                4:  /cgn2_6/ptodata/2/paa/US08_NEW_COMB.pep:*
                5:  /cgn2_6/ptodata/2/paa/US09_NEW_COMB.pep:*
                6:  /cgn2_6/ptodata/2/paa/US60_NEW_COMB.pep:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                                    SUMMARIES
                  %
Result          Query
  No.   Score   Match Length DB  ID                        Description
 -----------------------------------------------------------------------
    1      56   100.0      9  5  US-09-938-864-49           Sequence 49, Appl
    2      56   100.0      9  5  US-09-938-864-258          Sequence 258, App
    3      56   100.0    133  5  US-09-938-864-344          Sequence 344, App
    4      56   100.0    214  5  US-09-938-864-395          Sequence 395, App
    5      56   100.0    256  5  US-09-938-864-335          Sequence 335, App
    6      56   100.0    344  5  US-09-938-864-391          Sequence 391, App
    7      56   100.0    345  5  US-09-929-315-2            Sequence 2, Appli
    8      56   100.0    362  5  US-09-938-864-394          Sequence 394, App
    9      56   100.0    369  5  US-09-938-864-346          Sequence 346, App
   10      56   100.0    410  5  US-09-938-864-333          Sequence 333, App
   11      56   100.0    420  5  US-09-938-864-393          Sequence 393, App
   12      56   100.0    429  5  US-09-938-864-408          Sequence 408, App
   13      56   100.0    449  5  US-09-538-092-960          Sequence 960, App
   14      56   100.0    449  5  US-09-929-315-4            Sequence 4, Appli
   15      56   100.0    449  5  US-09-929-315-6            Sequence 6, Appli
   16      56   100.0    449  5  US-09-938-864-319          Sequence 319, App
   17      56   100.0    449  5  US-09-938-864-320          Sequence 320, App
   18      56   100.0    449  5  US-09-938-864-404          Sequence 404, App
   19      56   100.0    495  5  US-09-938-864-409          Sequence 409, App
   20      56   100.0    504  5  US-09-938-864-410          Sequence 410, App
   21      56   100.0    568  5  US-09-938-864-392          Sequence 392, App
   22      56   100.0    591  5  US-09-938-864-334          Sequence 334, App
   23      46    82.1      9  5  US-09-938-864-131          Sequence 131, App
   24      46    82.1      9  5  US-09-938-864-278          Sequence 278, App
   25      41    73.2      9  5  US-09-938-864-167          Sequence 167, App
   26      41    73.2    154  5  US-09-929-315-5            Sequence 5, Appli
   27      37    66.1      9  5  US-09-938-864-245          Sequence 245, App
```

```
   28      37    66.1      9  5  US-09-938-864-305          Sequence 305, App
   29      37    66.1    255  5  US-09-760-466-856          Sequence 856, App
   30      36    64.3    428  5  US-09-938-864-405          Sequence 405, App
   31      36    64.3    436  6  US-60-314-050-7185         Sequence 7185, Ap
   32      35    62.5    333  1  PCT-US01-08631-45853       Sequence 45853, A
   33      35    62.5    345  6  US-60-312-544-9009         Sequence 9009, Ap
   34      34    60.7     29  5  US-09-864-761-44716        Sequence 44716, A
   35      34    60.7     44  1  PCT-US01-18569-3949        Sequence 3949, Ap
   36      34    60.7     72  5  US-09-758-445-575          Sequence 575, App
   37      34    60.7     72  5  US-09-757-025-554          Sequence 554, App
   38      34    60.7    371  1  PCT-US01-26189-66          Sequence 66, Appl
   39      34    60.7    371  5  US-09-934-455-66           Sequence 66, Appl
   40      34    60.7    385  5  US-09-538-092-333          Sequence 333, App
   41      34    60.7    492  5  US-09-570-581A-932         Sequence 932, App
   42      34    60.7    645  1  PCT-US01-08631-58241       Sequence 58241, A
   43      33    58.9     36  5  US-09-950-082-896          Sequence 896, App
   44      33    58.9    125  5  US-09-834-366-25075        Sequence 25075, A
   45      33    58.9    158  1  PCT-US01-08656-8691        Sequence 8691, Ap


                                   ALIGNMENTS


RESULT   1
US-09-938-864-49
; Sequence 49, Application US/09938864
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Smithgall, Molly
;  APPLICANT:  Moulton, Gus
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Sleath, Paul R.
;  APPLICANT:  Mossman, Sally
;  APPLICANT:  Evans, Lawrence
;  APPLICANT:  Spies, A. Gregory
;  APPLICANT:  Boydston, Jeremy
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1
;  TITLE OF INVENTION:  SPECIFIC IMMUNOTHERAPY
;  FILE REFERENCE: 210121.465C5
;  CURRENT APPLICATION NUMBER: US/09/938,864
;  CURRENT FILING DATE:  2001-08-24
;  NUMBER OF SEQ ID NOS: 413
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 49
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Homo sapien
US-09-938-864-49


  Query Match             100.0%;  Score 56;  DB 5;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.9e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CMTWNQMNL 9
           |||||||||
Db       1 CMTWNQMNL 9


RESULT   2
US-09-938-864-258
; Sequence 258, Application US/09938864
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Smithgall, Molly
;  APPLICANT:  Moulton, Gus
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Sleath, Paul R.
;  APPLICANT:  Mossman, Sally
;  APPLICANT:  Evans, Lawrence
;  APPLICANT:  Spies, A. Gregory
```

```
;  TITLE OF INVENTION: METHODS FOR THE DIAGNOSIS AND THERAPY
;  TITLE OF INVENTION: OF CHLAMYDIA INFECTION
;  FILE REFERENCE: 210121.492
;  CURRENT APPLICATION NUMBER: US/09/523,419
;  CURRENT FILING DATE: 2000-03-10
;  NUMBER OF SEQ ID NOS: 326
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 258
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-523-419-258


  Query Match             100.0%;  Score 56;  DB 19;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.6e+06;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CMTWNQMNL 9
           |||||||||
Db       1 CMTWNQMNL 9

RESULT  13
US-09-625-963-2
; Sequence 2, Application US/09625963
; GENERAL INFORMATION:
;  APPLICANT:  Strauss, Hans Josef
;  APPLICANT:  Gao, Liquan
;  TITLE OF INVENTION:  Immunotherapeutic Methods Using Epitopes of WT-1 and
;  TITLE OF INVENTION:  GATA-1
;  FILE REFERENCE:  ICI 101
;  CURRENT APPLICATION NUMBER: US/09/625,963
;  CURRENT FILING DATE:  2000-07-26
;  PRIOR APPLICATION NUMBER:  PCT/GB99/03572
;  PRIOR FILING DATE:  1999-11-02
;  PRIOR APPLICATION NUMBER:  GB9823897.5
;  PRIOR FILING DATE:  1998-11-02
;  NUMBER OF SEQ ID NOS:  5
;  SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 2
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-625-963-2


  Query Match             100.0%;  Score 56;  DB 20;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.6e+06;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CMTWNQMNL 9
           |||||||||
Db       1 CMTWNQMNL 9

RESULT  14
US-09-679-339-49
; Sequence 49, Application US/09679339
; GENERAL INFORMATION:
;  APPLICANT:  Skeiky, Yasir
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Cheever, Martin
;  APPLICANT:  Reed, Steven G.
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR WT1
;  TITLE OF INVENTION:  SPECIFIC IMMUNOTHERAPY
;  FILE REFERENCE:  210121.533
;  CURRENT APPLICATION NUMBER: US/09/679,339
;  CURRENT FILING DATE:  2000-10-04
;  NUMBER OF SEQ ID NOS:  357
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 49
;   LENGTH: 9
```

```
;   TYPE: PRT
;   ORGANISM: Homo sapien
US-09-679-339-49


  Query Match             100.0%;  Score 56;  DB 20;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.6e+06;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CMTWNQMNL 9
           |||||||||
Db       1 CMTWNQMNL 9

RESULT  15
US-09-679-339-258
; Sequence 258, Application US/09679339
; GENERAL INFORMATION:
;  APPLICANT:  Skeiky, Yasir
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Cheever, Martin
;  APPLICANT:  Reed, Steven G.
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR WT1
;  TITLE OF INVENTION:  SPECIFIC IMMUNOTHERAPY
;  FILE REFERENCE:  210121.533
;  CURRENT APPLICATION NUMBER: US/09/679,339
;  CURRENT FILING DATE:  2000-10-04
;  NUMBER OF SEQ ID NOS:  357
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 258
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-679-339-258


  Query Match             100.0%;  Score 56;  DB 20;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.6e+06;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CMTWNQMNL 9
           |||||||||
Db       1 CMTWNQMNL 9


Search completed: October 23, 2001, 13:27:00
Job time: 413 sec
```

```
;  NUMBER OF SEQ ID NOS: 326
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 49
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Homo sapien
US-09-276-484-49


  Query Match             100.0%;  Score 56;  DB 16;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.6e+06;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CMTWNQMNL 9
           |||||||||
Db       1 CMTWNQMNL 9


RESULT   8
US-09-276-484-258
; Sequence 258, Application US/09276484
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
;  APPLICANT: Cheever, Martin A.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1
;  TITLE OF INVENTION: SPECIFIC IMMUNOTHERAPY
;  FILE REFERENCE: 210121.465C1
;  CURRENT APPLICATION NUMBER: US/09/276,484
;  CURRENT FILING DATE: 1999-03-25
;  NUMBER OF SEQ ID NOS: 326
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 258
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-276-484-258


  Query Match             100.0%;  Score 56;  DB 16;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.6e+06;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CMTWNQMNL 9
           |||||||||
Db       1 CMTWNQMNL 9


RESULT   9
US-09-276-484A-49
; Sequence 49, Application US/09276484A
; GENERAL INFORMATION:
;  APPLICANT:  Gaiger, Alexander
;  APPLICANT:  Cheever, Martin A.
;  TITLE OF INVENTION:  COMPOSITONS AND METHODS FOR WT1 SPECIFIC
;  TITLE OF INVENTION:  IMMUNOTHERAPY
;  FILE REFERENCE:  210121.465C1
;  CURRENT APPLICATION NUMBER: US/09/276,484A
;  CURRENT FILING DATE:  1999-03-25
;  PRIOR APPLICATION NUMBER:  US 09/164,223
;  PRIOR FILING DATE:  1999-09-30
;  NUMBER OF SEQ ID NOS:  326
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 49
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Homo sapien
US-09-276-484A-49


  Query Match             100.0%;  Score 56;  DB 16;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.6e+06;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy       1 CMTWNQMNL 9
           |||||||||
Db       1 CMTWNQMNL 9


RESULT  10
US-09-276-484A-258
; Sequence 258, Application US/09276484A
; GENERAL INFORMATION:
;  APPLICANT:  Gaiger, Alexander
;  APPLICANT:  Cheever, Martin A.
;  TITLE OF INVENTION:  COMPOSITONS AND METHODS FOR WT1 SPECIFIC
;  TITLE OF INVENTION:  IMMUNOTHERAPY
;  FILE REFERENCE:  210121.465C1
;  CURRENT APPLICATION NUMBER: US/09/276,484A
;  CURRENT FILING DATE:  1999-03-25
;  PRIOR APPLICATION NUMBER:  US 09/164,223
;  PRIOR FILING DATE:  1999-09-30
;  NUMBER OF SEQ ID NOS:  326
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 258
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-276-484A-258


  Query Match             100.0%;  Score 56;  DB 16;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.6e+06;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CMTWNQMNL 9
           |||||||||
Db       1 CMTWNQMNL 9


RESULT  11
US-09-523-419-49
; Sequence 49, Application US/09523419
; GENERAL INFORMATION:
;  APPLICANT: Hendrickson, Ronald C.
;  APPLICANT: Probst, Peter
;  APPLICANT: Reed, Steven G.
;  TITLE OF INVENTION: METHODS FOR THE DIAGNOSIS AND THERAPY
;  TITLE OF INVENTION: OF CHLAMYDIA INFECTION
;  FILE REFERENCE: 210121.492
;  CURRENT APPLICATION NUMBER: US/09/523,419
;  CURRENT FILING DATE: 2000-03-10
;  NUMBER OF SEQ ID NOS: 326
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 49
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Homo sapien
US-09-523-419-49


  Query Match             100.0%;  Score 56;  DB 19;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.6e+06;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CMTWNQMNL 9
           |||||||||
Db       1 CMTWNQMNL 9


RESULT  12
US-09-523-419-258
; Sequence 258, Application US/09523419
; GENERAL INFORMATION:
;  APPLICANT: Hendrickson, Ronald C.
;  APPLICANT: Probst, Peter
;  APPLICANT: Reed, Steven G.
```

```
;  APPLICANT: Cheever, Martin A.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1 SPECIFIC IMMUNOTHERAPY
;  FILE REFERENCE: 210121.465
;  CURRENT APPLICATION NUMBER: US/09/164,223
;  CURRENT FILING DATE: 1998-09-30
;  NUMBER OF SEQ ID NOS: 320
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 258
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-164-223-258


  Query Match             100.0%;  Score 56;  DB 15;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.6e+06;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 CMTWNQMNL 9
            |||||||||
Db        1 CMTWNQMNL 9

RESULT   3
US-09-164-223-49
; Sequence 49, Application US/09164223A
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
;  APPLICANT: Cheever, Martin A.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1 SPECIFIC IMMUNOTHERAPY
;  FILE REFERENCE: 210121.465
;  CURRENT APPLICATION NUMBER: US/09/164,223A
;  CURRENT FILING DATE: 1998-09-30
;  NUMBER OF SEQ ID NOS: 320
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 49
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-164-223-49


  Query Match             100.0%;  Score 56;  DB 15;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.6e+06;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 CMTWNQMNL 9
            |||||||||
Db        1 CMTWNQMNL 9

RESULT   4
US-09-164-223-258
; Sequence 258, Application US/09164223A
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
;  APPLICANT: Cheever, Martin A.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1 SPECIFIC IMMUNOTHERAPY
;  FILE REFERENCE: 210121.465
;  CURRENT APPLICATION NUMBER: US/09/164,223A
;  CURRENT FILING DATE: 1998-09-30
;  NUMBER OF SEQ ID NOS: 320
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 258
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-164-223-258


  Query Match             100.0%;  Score 56;  DB 15;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.6e+06;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy        1 CMTWNQMNL 9
            |||||||||
Db        1 CMTWNQMNL 9


RESULT   5
US-09-164-223-49
; Sequence 49, Application US/09164223B
; GENERAL INFORMATION:
;  APPLICANT:  Gaiger, Alexander
;  APPLICANT:  Cheever, Martin A.
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR WT1 SPECIFIC IMMUNOTHERAPY
;  FILE REFERENCE:  210121.465
;  CURRENT APPLICATION NUMBER: US/09/164,223B
;  CURRENT FILING DATE:  1998-09-30
;  NUMBER OF SEQ ID NOS:  320
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 49
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-164-223-49


  Query Match             100.0%;  Score 56;  DB 15;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.6e+06;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 CMTWNQMNL 9
            |||||||||
Db        1 CMTWNQMNL 9

RESULT   6
US-09-164-223-258
; Sequence 258, Application US/09164223B
; GENERAL INFORMATION:
;  APPLICANT:  Gaiger, Alexander
;  APPLICANT:  Cheever, Martin A.
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR WT1 SPECIFIC IMMUNOTHERAPY
;  FILE REFERENCE:  210121.465
;  CURRENT APPLICATION NUMBER: US/09/164,223B
;  CURRENT FILING DATE:  1998-09-30
;  NUMBER OF SEQ ID NOS:  320
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 258
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-164-223-258


  Query Match             100.0%;  Score 56;  DB 15;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.6e+06;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 CMTWNQMNL 9
            |||||||||
Db        1 CMTWNQMNL 9

RESULT   7
US-09-276-484-49
; Sequence 49, Application US/09276484
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
;  APPLICANT: Cheever, Martin A.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1
;  TITLE OF INVENTION: SPECIFIC IMMUNOTHERAPY
;  FILE REFERENCE: 210121.465C1
;  CURRENT APPLICATION NUMBER: US/09/276,484
;  CURRENT FILING DATE: 1999-03-25
```

```
                      GenCore version 4.5
              Copyright (c) 1993 - 2000  Compugen Ltd.

OM protein - protein search, using sw model

Run on:          October 23, 2001, 13:26:59 ; Search time 241.8 Seconds
                                             (without alignments)
                                             8.257 Million cell updates/sec

Title:           US-09-625-963-2
Perfect score:   56
Sequence:        1 CMTWNQMNL 9

Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched:        2803329 seqs, 221847457 residues

Total number of hits satisfying chosen parameters:      2803329

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       Pending_Patents_AA_Main:*
                1:  /cgn2_6/ptodata/2/paa/PCTUS_COMB.pep:*
                2:  /cgn2_6/ptodata/2/paa/US06_COMB.pep:*
                3:  /cgn2_6/ptodata/2/paa/US07_COMB.pep:*
                4:  /cgn2_6/ptodata/2/paa/US080_COMB.pep:*
                5:  /cgn2_6/ptodata/2/paa/US081_COMB.pep:*
                6:  /cgn2_6/ptodata/2/paa/US082_COMB.pep:*
                7:  /cgn2_6/ptodata/2/paa/US083_COMB.pep:*
                8:  /cgn2_6/ptodata/2/paa/US084_COMB.pep:*
                9:  /cgn2_6/ptodata/2/paa/US085_COMB.pep:*
                10:  /cgn2_6/ptodata/2/paa/US086_COMB.pep:*
                11:  /cgn2_6/ptodata/2/paa/US087_COMB.pep:*
                12:  /cgn2_6/ptodata/2/paa/US088_COMB.pep:*
                13:  /cgn2_6/ptodata/2/paa/US089_COMB.pep:*
                14:  /cgn2_6/ptodata/2/paa/US090_COMB.pep:*
                15:  /cgn2_6/ptodata/2/paa/US091_COMB.pep:*
                16:  /cgn2_6/ptodata/2/paa/US092_COMB.pep:*
                17:  /cgn2_6/ptodata/2/paa/US093_COMB.pep:*
                18:  /cgn2_6/ptodata/2/paa/US094_COMB.pep:*
                19:  /cgn2_6/ptodata/2/paa/US095_COMB.pep:*
                20:  /cgn2_6/ptodata/2/paa/US096_COMB.pep:*
                21:  /cgn2_6/ptodata/2/paa/US097_COMB.pep:*
                22:  /cgn2_6/ptodata/2/paa/US098_COMB.pep:*
                23:  /cgn2_6/ptodata/2/paa/US60_COMB.pep:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                                  SUMMARIES
                     %
Result               Query
  No.    Score   Match Length DB  ID                         Description
--------------------------------------------------------------------------------
    1      56    100.0       9  15  US-09-164-223-49           Sequence 49, Appl
    2      56    100.0       9  15  US-09-164-223-258          Sequence 258, App
    3      56    100.0       9  15  US-09-164-223-49           Sequence 49, Appl
    4      56    100.0       9  15  US-09-164-223-258          Sequence 258, App
    5      56    100.0       9  15  US-09-164-223-49           Sequence 49, Appl
    6      56    100.0       9  15  US-09-164-223-258          Sequence 258, App
    7      56    100.0       9  16  US-09-276-484-49           Sequence 49, Appl
    8      56    100.0       9  16  US-09-276-484-258          Sequence 258, App
    9      56    100.0       9  16  US-09-276-484A-49          Sequence 49, Appl
   10      56    100.0       9  16  US-09-276-484A-258         Sequence 258, App
```

```
   11      56    100.0       9  19  US-09-523-419-49           Sequence 49, Appl
   12      56    100.0       9  19  US-09-523-419-258          Sequence 258, App
   13      56    100.0       9  20  US-09-625-963-2            Sequence 2, Appli
   14      56    100.0       9  20  US-09-679-339-49           Sequence 49, Appl
   15      56    100.0       9  20  US-09-679-339-258          Sequence 258, App
   16      56    100.0       9  20  US-09-684-361-49           Sequence 49, Appl
   17      56    100.0       9  20  US-09-684-361-258          Sequence 258, App
   18      56    100.0       9  20  US-09-685-830-49           Sequence 49, Appl
   19      56    100.0       9  20  US-09-685-830-258          Sequence 258, App
   20      56    100.0       9  21  US-09-785-019-49           Sequence 49, Appl
   21      56    100.0       9  21  US-09-785-019-258          Sequence 258, App
   22      56    100.0       9  21  US-09-791-477-49           Sequence 49, Appl
   23      56    100.0       9  21  US-09-791-477-258          Sequence 258, App
   24      56    100.0     133  21  US-09-785-019-344          Sequence 344, App
   25      56    100.0     256  21  US-09-785-019-335          Sequence 335, App
   26      56    100.0     345  14  US-09-037-179A-2           Sequence 2, Appli
   27      56    100.0     345  14  US-09-037-179B-2           Sequence 2, Appli
   28      56    100.0     369  21  US-09-785-019-346          Sequence 346, App
   29      56    100.0     410  21  US-09-785-019-333          Sequence 333, App
   30      56    100.0     449  14  US-09-037-179A-4           Sequence 4, Appli
   31      56    100.0     449  14  US-09-037-179A-6           Sequence 6, Appli
   32      56    100.0     449  14  US-09-037-179B-4           Sequence 4, Appli
   33      56    100.0     449  14  US-09-037-179B-6           Sequence 6, Appli
   34      56    100.0     449  15  US-09-164-223-319          Sequence 319, App
   35      56    100.0     449  15  US-09-164-223-320          Sequence 320, App
   36      56    100.0     449  15  US-09-164-223-319          Sequence 319, App
   37      56    100.0     449  15  US-09-164-223-320          Sequence 320, App
   38      56    100.0     449  15  US-09-164-223-319          Sequence 319, App
   39      56    100.0     449  15  US-09-164-223-320          Sequence 320, App
   40      56    100.0     449  16  US-09-276-484-319          Sequence 319, App
   41      56    100.0     449  16  US-09-276-484-320          Sequence 320, App
   42      56    100.0     449  16  US-09-276-484A-319         Sequence 319, App
   43      56    100.0     449  16  US-09-276-484A-320         Sequence 320, App
   44      56    100.0     449  19  US-09-523-419-319          Sequence 319, App
   45      56    100.0     449  21  US-09-791-477-320          Sequence 320, App

                                 ALIGNMENTS


RESULT   1
US-09-164-223-49
; Sequence 49, Application US/09164223
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
;  APPLICANT: Cheever, Martin A.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1 SPECIFIC IMMUNOTHERAPY
;  FILE REFERENCE: 210121.465
;  CURRENT APPLICATION NUMBER: US/09/164,223
;  CURRENT FILING DATE: 1998-09-30
;  NUMBER OF SEQ ID NOS: 320
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 49
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-164-223-49


  Query Match             100.0%;  Score 56;  DB 15;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.6e+06;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CMTWNQMNL 9
           |||||||||
Db       1 CMTWNQMNL 9


RESULT   2
US-09-164-223-258
; Sequence 258, Application US/09164223
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
```

```
XX
OS   Homo sapiens.
XX
PN   WO200018795-A2.
XX
PD   06-APR-2000.
XX
PF   30-SEP-1999;   99WO-US22819.
XX
PR   30-SEP-1998;   98US-0164223.
PR   25-MAR-1999;   99US-0276484.
XX
PA   (CORI-) CORIXA CORP.
PA   (GAIG/) GAIGER A.
XX
PI   Gaiger A,  Cheever M;
XX
DR   WPI; 2000-293107/25.
XX
PT   Novel polypeptides comprising an immunogenic portion of a native WT1
PT   polypeptide, useful for inhibiting the development of malignant
PT   diseases associated with WT1 expression e.g. leukemia or cancer  -
XX
PS   Claim 4; Page 168; 193pp; English.
XX
CC   The present invention describes polypeptides (I) comprising an
CC   immunogenic portion of a native Wilm's Tumour gene product polypeptide,
CC   WT1, (or variants of the immunogenic portion retaining the ability to
CC   react with WT1-specific antisera and/or T-cell lines or clones) and
CC   comprising 16 consecutive amino acids (aa) or less of a native WT1
CC   polypeptide. The polypeptides are useful therapeutically and to
CC   manufacture medicaments for enhancing/inducing an immune response in
CC   patients. The polypeptides, mimetics or polynucleotides can be included
CC   with a carrier/excipient in pharmaceutical compositions or with a
CC   non-specific immune response enhancer (e.g. an adjuvant or enhancer
CC   preferentially enhancing a T cell response) in vaccines. Pharmaceutical
CC   compositions and vaccines can be administered to human patients to
CC   enhance or induce an immune response specific for WT1 or a cell
CC   expressing WT1, useful to inhibit the development of malignant diseases
CC   associated with WT1 expression, e.g. leukemia (especially acute/chronic
CC   myeloid leukemia or acute lymphocytic leukemia) or cancer (especially
CC   breast, lung, thyroid or gastrointestinal cancer, or a melanoma).
CC   AAY98501 to AAY98811 represent polypeptide sequences, and AAA13848 to
CC   AAA13862 represent PCR primers, used in the exemplification of the
CC   present invention.
XX
SQ   Sequence   9 AA;

  Query Match             73.2%;  Score 41;  DB 21;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 3.4e+05;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CMTWNQ 6
           ||||||
Db       4 cmtwnq 9


Search completed: October 23, 2001, 13:21:02
Job time: 55 sec
```

```
XX
KW   WT1; immunotherapy; immunogenic; malignant disease; cancer; leukaemia;
KW   metastatic disease; mouse; human; Wilm's tumour; immune response;
KW   vaccine.
XX
OS   Homo sapiens.
XX
PN   WO200018795-A2.
XX
PD   06-APR-2000.
XX
PF   30-SEP-1999;   99WO-US22819.
XX
PR   30-SEP-1998;   98US-0164223.
PR   25-MAR-1999;   99US-0276484.
XX
PA   (CORI-) CORIXA CORP.
PA   (GAIG/) GAIGER A.
XX
PI   Gaiger A,  Cheever M;
XX
DR   WPI; 2000-293107/25.
XX
PT   Novel polypeptides comprising an immunogenic portion of a native WT1
PT   polypeptide, useful for inhibiting the development of malignant
PT   diseases associated with WT1 expression e.g. leukemia or cancer  -
XX
PS   Claim 4; Page 163; 193pp; English.
XX
CC   The present invention describes polypeptides (I) comprising an
CC   immunogenic portion of a native Wilm's Tumour gene product polypeptide,
CC   WT1, (or variants of the immunogenic portion retaining the ability to
CC   react with WT1-specific antisera and/or T-cell lines or clones) and
CC   comprising 16 consecutive amino acids (aa) or less of a native WT1
CC   polypeptide. The polypeptides are useful therapeutically and to
CC   manufacture medicaments for enhancing/inducing an immune response in
CC   patients. The polypeptides, mimetics or polynucleotides can be included
CC   with a carrier/excipient in pharmaceutical compositions or with a
CC   non-specific immune response enhancer (e.g. an adjuvant or enhancer
CC   preferentially enhancing a T cell response) in vaccines. Pharmaceutical
CC   compositions and vaccines can be administered to human patients to
CC   enhance or induce an immune response specific for WT1 or a cell
CC   expressing WT1, useful to inhibit the development of malignant diseases
CC   associated with WT1 expression, e.g. leukemia (especially acute/chronic
CC   myeloid leukemia or acute lymphocytic leukemia) or cancer (especially
CC   breast, lung, thyroid or gastrointestinal cancer, or a melanoma).
CC   AAY98501 to AAY98811 represent polypeptide sequences, and AAA13848 to
CC   AAA13862 represent PCR primers, used in the exemplification of the
CC   present invention.
XX
SQ   Sequence   9 AA;


  Query Match             82.1%;  Score 46;  DB 21;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 3.4e+05;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CMTWNQM 7
           |||||||
Db       3 cmtwnqm 9


RESULT  14
AAY98763
ID   AAY98763 standard; Peptide; 9 AA.
XX
AC   AAY98763;
XX
DT   31-JUL-2000  (first entry)
XX
DE   WT1 derived immunogenic peptide SEQ ID NO:278.
XX
KW   WT1; immunotherapy; immunogenic; malignant disease; cancer; leukaemia;
```

```
KW   metastatic disease; mouse; human; Wilm's tumour; immune response;
KW   vaccine.
XX
OS   Mus musculus.
XX
PN   WO200018795-A2.
XX
PD   06-APR-2000.
XX
PF   30-SEP-1999;   99WO-US22819.
XX
PR   30-SEP-1998;   98US-0164223.
PR   25-MAR-1999;   99US-0276484.
XX
PA   (CORI-) CORIXA CORP.
PA   (GAIG/) GAIGER A.
XX
PI   Gaiger A,  Cheever M;
XX
DR   WPI; 2000-293107/25.
XX
PT   Novel polypeptides comprising an immunogenic portion of a native WT1
PT   polypeptide, useful for inhibiting the development of malignant
PT   diseases associated with WT1 expression e.g. leukemia or cancer  -
XX
PS   Claim 4; Page 184; 193pp; English.
XX
CC   The present invention describes polypeptides (I) comprising an
CC   immunogenic portion of a native Wilm's Tumour gene product polypeptide,
CC   WT1, (or variants of the immunogenic portion retaining the ability to
CC   react with WT1-specific antisera and/or T-cell lines or clones) and
CC   comprising 16 consecutive amino acids (aa) or less of a native WT1
CC   polypeptide. The polypeptides are useful therapeutically and to
CC   manufacture medicaments for enhancing/inducing an immune response in
CC   patients. The polypeptides, mimetics or polynucleotides can be included
CC   with a carrier/excipient in pharmaceutical compositions or with a
CC   non-specific immune response enhancer (e.g. an adjuvant or enhancer
CC   preferentially enhancing a T cell response) in vaccines. Pharmaceutical
CC   compositions and vaccines can be administered to human patients to
CC   enhance or induce an immune response specific for WT1 or a cell
CC   expressing WT1, useful to inhibit the development of malignant diseases
CC   associated with WT1 expression, e.g. leukemia (especially acute/chronic
CC   myeloid leukemia or acute lymphocytic leukemia) or cancer (especially
CC   breast, lung, thyroid or gastrointestinal cancer, or a melanoma).
CC   AAY98501 to AAY98811 represent polypeptide sequences, and AAA13848 to
CC   AAA13862 represent PCR primers, used in the exemplification of the
CC   present invention.
XX
SQ   Sequence   9 AA;

  Query Match             82.1%;  Score 46;  DB 21;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 3.4e+05;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CMTWNQM 7
           |||||||
Db       3 cmtwnqm 9


RESULT  15
AAY98652
ID   AAY98652 standard; Peptide; 9 AA.
XX
AC   AAY98652;
XX
DT   31-JUL-2000  (first entry)
XX
DE   WT1 derived immunogenic peptide SEQ ID NO:167.
XX
KW   WT1; immunotherapy; immunogenic; malignant disease; cancer; leukaemia;
KW   metastatic disease; mouse; human; Wilm's tumour; immune response;
KW   vaccine.
```

```
CC   enhance or induce an immune response specific for WT1 or a cell
CC   expressing WT1, useful to inhibit the development of malignant diseases
CC   associated with WT1 expression, e.g. leukemia (especially acute/chronic
CC   myeloid leukemia or acute lymphocytic leukemia) or cancer (especially
CC   breast, lung, thyroid or gastrointestinal cancer, or a melanoma).
CC   AAY98501 to AAY98811 represent polypeptide sequences, and AAA13848 to
CC   AAA13862 represent PCR primers, used in the exemplification of the
CC   present invention.
XX
SQ   Sequence   449 AA;

  Query Match             100.0%;  Score 56;  DB 21;  Length 449;
  Best Local Similarity   100.0%;  Pred. No. 0.14;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CMTWNQMNL 9
             |||||||||
Db     235 cmtwnqmnl 243


RESULT  11
AAY80196
ID   AAY80196 standard; protein; 449 AA.
XX
AC   AAY80196;
XX
DT   24-MAY-2000  (first entry)
XX
DE   Mouse Wilms' tumour suppressor gene WT1 product SEQ ID NO:1.
XX
KW   Wilm's tumour suppressor gene; WT1; cancer; antigen; vaccine; MHC;
KW   major histocompatibility complex; leukaemia; tumour; antitumour.
XX
OS   Mus sp.
XX
PN   WO200006602-A1.
XX
PD   10-FEB-2000.
XX
PF   30-JUL-1999;   99WO-JP04130.
XX
PR   31-JUL-1998;   98JP-0218093.
XX
PA   (SUGI/) SUGIYAMA H.
XX
PI   Sugiyama H,  Oka Y;
XX
DR   WPI; 2000-195264/17.
XX
PT   Cancer antigens based on Wilm's tumor suppressor gene WT1 product or
PT   peptide derivatives, for cancer vaccines in treating leukemia and solid
PT   tumors e.g. stomach cancer, skin cancer, lung cancer and breast cancer
XX
PS   Claim 2; Page 38-40; 48pp; Japanese.
XX
CC   The present invention describes a cancer antigen containing the active
CC   component of Wilms' tumour suppressor gene WT1 product, or partial
CC   peptides, for cancer vaccines in treating leukaemia and solid tumours.
CC   The cancer antigens are useful for cancer vaccines in treating
CC   leukaemia, bone-marrow abnormal formation syndrome, malignant lymphoma,
CC   multiple myeloma, stomach cancer, cancer of the large intestine, lung
CC   cancer, breast cancer, blastoma, liver cancer, skin cancer, bladder
CC   cancer, prostate cancer, uterus cancer, cervical cancer, or ovary
CC   cancer. The present sequence represents the mouse Wilms' tumour
CC   suppressor gene WT1 product.
XX
SQ   Sequence   449 AA;

  Query Match             100.0%;  Score 56;  DB 21;  Length 449;
  Best Local Similarity   100.0%;  Pred. No. 0.14;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy       1 CMTWNQMNL 9
             |||||||||
Db     235 cmtwnqmnl 243


RESULT  12
AAY80197
ID   AAY80197 standard; protein; 449 AA.
XX
AC   AAY80197;
XX
DT   24-MAY-2000  (first entry)
XX
DE   Human Wilms' tumour suppressor gene WT1 product SEQ ID NO:2.
XX
KW   Wilm's tumour suppressor gene; WT1; cancer; antigen; vaccine; MHC;
KW   major histocompatibility complex; leukaemia; tumour; antitumour.
XX
OS   Homo sapiens.
XX
PN   WO200006602-A1.
XX
PD   10-FEB-2000.
XX
PF   30-JUL-1999;   99WO-JP04130.
XX
PR   31-JUL-1998;   98JP-0218093.
XX
PA   (SUGI/) SUGIYAMA H.
XX
PI   Sugiyama H,  Oka Y;
XX
DR   WPI; 2000-195264/17.
XX
PT   Cancer antigens based on Wilm's tumor suppressor gene WT1 product or
PT   peptide derivatives, for cancer vaccines in treating leukemia and solid
PT   tumors e.g. stomach cancer, skin cancer, lung cancer and breast cancer
XX
PS   Claim 2; Page 40-42; 48pp; Japanese.
XX
CC   The present invention describes a cancer antigen containing the active
CC   component of Wilms' tumour suppressor gene WT1 product, or partial
CC   peptides, for cancer vaccines in treating leukaemia and solid tumours.
CC   The cancer antigens are useful for cancer vaccines in treating
CC   leukaemia, bone-marrow abnormal formation syndrome, malignant lymphoma,
CC   multiple myeloma, stomach cancer, cancer of the large intestine, lung
CC   cancer, breast cancer, blastoma, liver cancer, skin cancer, bladder
CC   cancer, prostate cancer, uterus cancer, cervical cancer, or ovary
CC   cancer. The present sequence represents the human Wilms' tumour
CC   suppressor gene WT1 product.
XX
SQ   Sequence   449 AA;

  Query Match             100.0%;  Score 56;  DB 21;  Length 449;
  Best Local Similarity   100.0%;  Pred. No. 0.14;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CMTWNQMNL 9
             |||||||||
Db     235 cmtwnqmnl 243


RESULT  13
AAY98616
ID   AAY98616 standard; Peptide; 9 AA.
XX
AC   AAY98616;
XX
DT   31-JUL-2000  (first entry)
XX
DE   WT1 derived immunogenic peptide SEQ ID NO:131.
```

```
PT   and antibody production
XX
PS   Disclosure; Columns 39-42; 42pp; English.
XX
CC   This is a Wilms' tumour (WT1) polypeptide. The Wilms' tumour gene
CC   (WT1) is associated with 11p3 locus on the human chromosome. The
CC   polypeptide or immunogenic fragments of the polypeptide can be used to
CC   treat cancerous or precancerous conditions (especially Wilms' tumour), or
CC   to raise antibodies for diagnostic use. The product allow detection of
CC   risk of development of Wilms' tumour, e.g. in diseases such as WAGR and
CC   Denys-Drash syndrome, to be assessed prior to current methods of
CC   detection.
XX
SQ   Sequence   449 AA;


  Query Match            100.0%;  Score 56;  DB 19;  Length 449;
  Best Local Similarity  100.0%;  Pred. No. 0.14;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CMTWNQMNL 9
           |||||||||
Db     235 cmtwnqmnl 243


RESULT   9
AAY98804
ID   AAY98804 standard; Protein; 449 AA.
XX
AC   AAY98804;
XX
DT   31-JUL-2000  (first entry)
XX
DE   Human WT1 protein sequence SEQ ID NO:319.
XX
KW   WT1; immunotherapy; immunogenic; malignant disease; cancer; leukaemia;
KW   metastatic disease; mouse; human; Wilm's tumour; immune response;
KW   vaccine.
XX
OS   Homo sapiens.
XX
PN   WO200018795-A2.
XX
PD   06-APR-2000.
XX
PF   30-SEP-1999;   99WO-US22819.
XX
PR   30-SEP-1998;   98US-0164223.
PR   25-MAR-1999;   99US-0276484.
XX
PA   (CORI-) CORIXA CORP.
PA   (GAIG/) GAIGER A.
XX
PI   Gaiger A,  Cheever M;
XX
DR   WPI; 2000-293107/25.
XX
PT   Novel polypeptides comprising an immunogenic portion of a native WT1
PT   polypeptide, useful for inhibiting the development of malignant
PT   diseases associated with WT1 expression e.g. leukemia or cancer  -
XX
PS   Disclosure; Page 190-191; 193pp; English.
XX
CC   The present invention describes polypeptides (I) comprising an
CC   immunogenic portion of a native Wilm's Tumour gene product polypeptide,
CC   WT1, (or variants of the immunogenic portion retaining the ability to
CC   react with WT1-specific antisera and/or T-cell lines or clones) and
CC   comprising 16 consecutive amino acids (aa) or less of a native WT1
CC   polypeptide. The polypeptides are useful therapeutically and to
CC   manufacture medicaments for enhancing/inducing an immune response in
CC   patients. The polypeptides, mimetics or polynucleotides can be included
CC   with a carrier/excipient in pharmaceutical compositions or with a
CC   non-specific immune response enhancer (e.g. an adjuvant or enhancer
```

```
CC   preferentially enhancing a T cell response) in vaccines. Pharmaceutical
CC   compositions and vaccines can be administered to human patients to
CC   enhance or induce an immune response specific for WT1 or a cell
CC   expressing WT1, useful to inhibit the development of malignant diseases
CC   associated with WT1 expression, e.g. leukemia (especially acute/chronic
CC   myeloid leukemia or acute lymphocytic leukemia) or cancer (especially
CC   breast, lung, thyroid or gastrointestinal cancer, or a melanoma).
CC   AAY98501 to AAY98811 represent polypeptide sequences, and AAA13848 to
CC   AAA13862 represent PCR primers, used in the exemplification of the
CC   present invention.
XX
SQ   Sequence   449 AA;

  Query Match            100.0%;  Score 56;  DB 21;  Length 449;
  Best Local Similarity  100.0%;  Pred. No. 0.14;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CMTWNQMNL 9
           |||||||||
Db     235 cmtwnqmnl 243


RESULT  10
AAY98805
ID   AAY98805 standard; Protein; 449 AA.
XX
AC   AAY98805;
XX
DT   31-JUL-2000  (first entry)
XX
DE   Mouse WT1 protein sequence SEQ ID NO:320.
XX
KW   WT1; immunotherapy; immunogenic; malignant disease; cancer; leukaemia;
KW   metastatic disease; mouse; human; Wilm's tumour; immune response;
KW   vaccine.
XX
OS   Mus musculus.
XX
PN   WO200018795-A2.
XX
PD   06-APR-2000.
XX
PF   30-SEP-1999;   99WO-US22819.
XX
PR   30-SEP-1998;   98US-0164223.
PR   25-MAR-1999;   99US-0276484.
XX
PA   (CORI-) CORIXA CORP.
PA   (GAIG/) GAIGER A.
XX
PI   Gaiger A,  Cheever M;
XX
DR   WPI; 2000-293107/25.
XX
PT   Novel polypeptides comprising an immunogenic portion of a native WT1
PT   polypeptide, useful for inhibiting the development of malignant
PT   diseases associated with WT1 expression e.g. leukemia or cancer  -
XX
PS   Disclosure; Page 191-192; 193pp; English.
XX
CC   The present invention describes polypeptides (I) comprising an
CC   immunogenic portion of a native Wilm's Tumour gene product polypeptide,
CC   WT1, (or variants of the immunogenic portion retaining the ability to
CC   react with WT1-specific antisera and/or T-cell lines or clones) and
CC   comprising 16 consecutive amino acids (aa) or less of a native WT1
CC   polypeptide. The polypeptides are useful therapeutically and to
CC   manufacture medicaments for enhancing/inducing an immune response in
CC   patients. The polypeptides, mimetics or polynucleotides can be included
CC   with a carrier/excipient in pharmaceutical compositions or with a
CC   non-specific immune response enhancer (e.g. an adjuvant or enhancer
CC   preferentially enhancing a T cell response) in vaccines. Pharmaceutical
CC   compositions and vaccines can be administered to human patients to
```

```
XX
AC   AAW47173;
XX
DT   01-JUN-1998  (first entry)
XX
DE   Wilms' tumour polypeptide (WT33).
XX
KW   Wilms' tumour; WT33; cancer treatment; antibody production; WAGR;
KW   Denys-drash syndrome; WT1.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Region          3..156
FT                   /note= "proline and glutamine rich region"
XX
PN   US5726288-A.
XX
PD   10-MAR-1998.
XX
PF   02-AUG-1993;   93US-0102942.
XX
PR   02-AUG-1993;   93US-0102942.
PR   13-NOV-1989;   89US-0435780.
PR   13-NOV-1990;   90US-0614161.
XX
PA   (MASI ) MASSACHUSETTS INST TECHNOLOGY.
XX
PI   Bruening W,  Buckler AJ,  Call KM,  Darveau A,  Glaser TM;
PI   Haber DA,  Housman DE,  Ito CY,  Pelletier J,  Rose EA;
XX
DR   WPI; 1998-192828/17.
DR   N-PSDB; AAV17060.
XX
PT   New Wilms' tumour polypeptide - useful for, e.g. cancer treatment
PT   and antibody production
XX
PS   Claim 2; Fig 3; 42pp; English.
XX
CC   This is a new Wilms' tumour (WT33) polypeptide. The Wilms' tumour gene
CC   (WT1) is associated with 11p3 locus on the human chromosome. This
CC   polypeptide has a region rich in proline and glutamine (AAW47175)
CC   indicating that it has a role in transcription regulation. The
CC   polypeptide or immunogenic fragments of the polypeptide can be used to
CC   treat cancerous or precancerous conditions (especially Wilms' tumour),
CC   or to raise antibodies for diagnostic use. The product allow detection
CC   of risk of development of Wilms' tumour, e.g. in diseases such as WAGR
CC   and Denys-Drash syndrome, to be assessed prior to current methods of
CC   detection.
XX
SQ   Sequence   345 AA;

  Query Match             100.0%;  Score 56;  DB 19;  Length 345;
  Best Local Similarity   100.0%;  Pred. No. 0.11;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 CMTWNQMNL 9
          |||||||||
Db    151 cmtwnqmnl 159


RESULT   7
AAR85065
ID   AAR85065 standard; Protein; 429 AA.
XX
AC   AAR85065;
XX
DT   02-FEB-1996  (first entry)
XX
DE   Wilms' tumour WT1 antigen.
XX
KW   Wilms' tumour; WT1; monoclonal antibody; leukaemia.
```

```
XX
OS   Homo sapiens.
XX
PN   WO9529995-A1.
XX
PD   09-NOV-1995.
XX
PF   25-APR-1995;   95WO-US05523.
XX
PR   28-APR-1994;   94US-0234783.
XX
PA   (WIST-) WISTAR INST ANATOMY & BIOLOGY.
XX
PI   Herlyn M,  Morris J,  Rauscher FJ,  Rodeck U;
XX
DR   WPI; 1995-393072/50.
DR   N-PSDB; AAR85065.
XX
PT   Monoclonal antibodies specific for Wilms' tumour protein antigen WT1
PT   - useful for detecting, monitoring and diagnosing malignancies
PT   characterised by expression of WT1 protein, e.g. leukaemia
XX
PS   Claim 1; Page 42-43; 54pp; English.
XX
CC   The human Wilm's tumour antigen WT1 (sequence given in AAR85065) was
CC   used as the basis for the design of a synthetic WT1 sequence (see
CC   AAR85066) used to raise monoclonal antibodies specific for the antigen.
XX
SQ   Sequence   429 AA;

  Query Match             100.0%;  Score 56;  DB 16;  Length 429;
  Best Local Similarity   100.0%;  Pred. No. 0.14;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 CMTWNQMNL 9
          |||||||||
Db    235 cmtwnqmnl 243


RESULT   8
AAW47176
ID   AAW47176 standard; Protein; 449 AA.
XX
AC   AAW47176;
XX
DT   01-JUN-1998  (first entry)
XX
DE   Wilms' tumour polypeptide (WT1).
XX
KW   Wilms' tumour; WT33; cancer treatment; antibody production; WAGR;
KW   Denys-drash syndrome; WT1.
XX
OS   Homo sapiens.
XX
PN   US5726288-A.
XX
PD   10-MAR-1998.
XX
PF   02-AUG-1993;   93US-0102942.
XX
PR   02-AUG-1993;   93US-0102942.
PR   13-NOV-1989;   89US-0435780.
PR   13-NOV-1990;   90US-0614161.
XX
PA   (MASI ) MASSACHUSETTS INST TECHNOLOGY.
XX
PI   Bruening W,  Buckler AJ,  Call KM,  Darveau A,  Glaser TM;
PI   Haber DA,  Housman DE,  Ito CY,  Pelletier J,  Rose EA;
XX
DR   WPI; 1998-192828/17.
XX
PT   New Wilms' tumour polypeptide - useful for, e.g. cancer treatment
```

```
CC   myeloid leukemia or acute lymphocytic leukemia) or cancer (especially
CC   breast, lung, thyroid or gastrointestinal cancer, or a melanoma).
CC   AAY98501 to AAY98811 represent polypeptide sequences, and AAA13848 to
CC   AAA13862 represent PCR primers, used in the exemplification of the
CC   present invention.
XX
SQ   Sequence   9 AA;


  Query Match              100.0%;  Score 56;  DB 21;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 3.4e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 CMTWNQMNL 9
            |||||||||
Db        1 cmtwnqmnl 9


RESULT   4
AAY80202
ID   AAY80202 standard; Peptide; 9 AA.
XX
AC   AAY80202;
XX
DT   24-MAY-2000  (first entry)
XX
DE   Human Wilms' tumour suppressor gene WT1 product peptide SEQ ID NO:7.
XX
KW   Wilm's tumour suppressor gene; WT1; cancer; antigen; vaccine; MHC;
KW   major histocompatibility complex; leukaemia; tumour; antitumour.
XX
OS   Homo sapiens.
XX
PN   WO200006602-A1.
XX
PD   10-FEB-2000.
XX
PF   30-JUL-1999;   99WO-JP04130.
XX
PR   31-JUL-1998;   98JP-0218093.
XX
PA   (SUGI/) SUGIYAMA H.
XX
PI   Sugiyama H,  Oka Y;
XX
DR   WPI; 2000-195264/17.
XX
PT   Cancer antigens based on Wilm's tumor suppressor gene WT1 product or
PT   peptide derivatives, for cancer vaccines in treating leukemia and solid
PT   tumors e.g. stomach cancer, skin cancer, lung cancer and breast cancer
XX
PS   Claim 5; Page 18; 48pp; Japanese.
XX
CC   The present invention describes a cancer antigen containing the active
CC   component of Wilms' tumour suppressor gene WT1 product, or partial
CC   peptides, for cancer vaccines in treating leukaemia and solid tumours.
CC   The cancer antigens are useful for cancer vaccines in treating
CC   leukaemia, bone-marrow abnormal formation syndrome, malignant lymphoma,
CC   multiple myeloma, stomach cancer, cancer of the large intestine, lung
CC   cancer, breast cancer, blastoma, liver cancer, skin cancer, bladder
CC   cancer, prostate cancer, uterus cancer, cervical cancer, or ovary
CC   cancer. The present sequence represents a peptide from the human
CC   Wilms' tumour suppressor gene WT1 product.
XX
SQ   Sequence   9 AA;


  Query Match              100.0%;  Score 56;  DB 21;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 3.4e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 CMTWNQMNL 9
            |||||||||
```

```
Db        1 cmtwnqmnl 9


RESULT   5
AAR12240
ID   AAR12240 standard; Protein; 345 AA.
XX
AC   AAR12240;
XX
DT   15-AUG-1991  (first entry)
XX
DE   Wilm's tumour gene prod.
XX
KW   Wilm's tumour; monoclonal antibodies.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Domain          3..156
FT                   /label= proline/glutamine rich region
FT   Domain          216..226
FT                   /label= zinc finger domain 1
FT   Domain          227..246
FT                   /label= zinc finger domain 2
FT   Domain          247..276
FT                   /label= zinc finger domain 3
FT   Domain          277..304
FT                   /label= zinc finger domain 4
XX
PN   WO9107509-A.
XX
PD   30-MAY-1991.
XX
PF   13-NOV-1990;   90WO-US06629.
XX
PR   13-NOV-1989;   89US-0435780.
XX
PA   (MASI ) MASSACHUSETTS INST TECH.
XX
PI   Call KM,  Glaser TM,  Ito CY,  Buckler AJ,  Pelletier J;
PI   Haber DA,  Rose EA,  Housman DE;
XX
DR   WPI; 1991-178117/24.
DR   N-PSDB; AAQ12020.
XX
PT   Nucleotide sequence of human chromosome 11 band 13 Wilm's tumour
PT   locus - used to diagnose, quantify and treat Wilm's tumours
XX
PS   Disclosure; fig 3; 67pp; English.
XX
CC   This polypeptide is encoded by the Wilm's tumour (WT) gene which
CC   maps to the 11p13 locus of human chromosome 11.  It contains 4
CC   zinc finger binding domains and a proline/glutamine rich region
CC   and hence is thought to be involved in transcription regulation.
CC   Using this WT protein and the DNA, probes and antibodies can be
CC   developed for diagnosing WT.  Lesions similar to WT may also
CC   be detected.
XX
SQ   Sequence   345 AA;


  Query Match              100.0%;  Score 56;  DB 12;  Length 345;
  Best Local Similarity   100.0%;  Pred. No. 0.11;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 CMTWNQMNL 9
              |||||||||
Db      151 cmtwnqmnl 159


RESULT   6
AAW47173
ID   AAW47173 standard; Protein; 345 AA.
```

```
CC   The present sequence is peptide epitope WT235-43, produced by WT1
CC   expressing cells and found at residues 235-243 of the WT1 protein, which
CC   is recognised by cytotoxic T lymphocytes. WT1 is aberrantly expressed in
CC   leukaemias, breast cancer, melanoma and ovarian cancer. The peptide can
CC   be used as a vaccine to stimulate the elimination, by cytotoxic T
CC   lymphocytes, of cancer cells aberrantly expressing WT1. In addition, the
CC   nucleic acid encoding the peptide may also be used in the same manner.
CC   Alternatively, the peptide may be used in vitro to produce activated
CC   cytotoxic T lymphocytes.
XX
SQ   Sequence   9 AA;

  Query Match             100.0%;  Score 56;  DB 21;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 3.4e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CMTWNQMNL 9
              |||||||||
Db          1 cmtwnqmnl 9


RESULT   2
AAY98534
ID   AAY98534 standard; Peptide; 9 AA.
XX
AC   AAY98534;
XX
DT   31-JUL-2000  (first entry)
XX
DE   WT1 derived immunogenic peptide SEQ ID NO:49.
XX
KW   WT1; immunotherapy; immunogenic; malignant disease; cancer; leukaemia;
KW   metastatic disease; mouse; human; Wilm's tumour; immune response;
KW   vaccine.
XX
OS   Homo sapiens.
XX
PN   WO200018795-A2.
XX
PD   06-APR-2000.
XX
PF   30-SEP-1999;   99WO-US22819.
XX
PR   30-SEP-1998;   98US-0164223.
PR   25-MAR-1999;   99US-0276484.
XX
PA   (CORI-) CORIXA CORP.
PA   (GAIG/) GAIGER A.
XX
PI   Gaiger A,  Cheever M;
XX
DR   WPI; 2000-293107/25.
XX
PT   Novel polypeptides comprising an immunogenic portion of a native WT1
PT   polypeptide, useful for inhibiting the development of malignant
PT   diseases associated with WT1 expression e.g. leukemia or cancer  -
XX
PS   Claim 4; Page 151; 193pp; English.
XX
CC   The present invention describes polypeptides (I) comprising an
CC   immunogenic portion of a native Wilm's Tumour gene product polypeptide,
CC   WT1, (or variants of the immunogenic portion retaining the ability to
CC   react with WT1-specific antisera and/or T-cell lines or clones) and
CC   comprising 16 consecutive amino acids (aa) or less of a native WT1
CC   polypeptide. The polypeptides are useful therapeutically and to
CC   manufacture medicaments for enhancing/inducing an immune response in
CC   patients. The polypeptides, mimetics or polynucleotides can be included
CC   with a carrier/excipient in pharmaceutical compositions or with a
CC   non-specific immune response enhancer (e.g. an adjuvant or enhancer
CC   preferentially enhancing a T cell response) in vaccines. Pharmaceutical
CC   compositions and vaccines can be administered to human patients to
CC   enhance or induce an immune response specific for WT1 or a cell
```

```
CC   expressing WT1, useful to inhibit the development of malignant diseases
CC   associated with WT1 expression, e.g. leukemia (especially acute/chronic
CC   myeloid leukemia or acute lymphocytic leukemia) or cancer (especially
CC   breast, lung, thyroid or gastrointestinal cancer, or a melanoma).
CC   AAY98501 to AAY98811 represent polypeptide sequences, and AAA13848 to
CC   AAA13862 represent PCR primers, used in the exemplification of the
CC   present invention.
XX
SQ   Sequence   9 AA;

  Query Match             100.0%;  Score 56;  DB 21;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 3.4e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CMTWNQMNL 9
              |||||||||
Db          1 cmtwnqmnl 9


RESULT   3
AAY98743
ID   AAY98743 standard; Peptide; 9 AA.
XX
AC   AAY98743;
XX
DT   31-JUL-2000  (first entry)
XX
DE   WT1 derived immunogenic peptide SEQ ID NO:258.
XX
KW   WT1; immunotherapy; immunogenic; malignant disease; cancer; leukaemia;
KW   metastatic disease; mouse; human; Wilm's tumour; immune response;
KW   vaccine.
XX
OS   Mus musculus.
XX
PN   WO200018795-A2.
XX
PD   06-APR-2000.
XX
PF   30-SEP-1999;   99WO-US22819.
XX
PR   30-SEP-1998;   98US-0164223.
PR   25-MAR-1999;   99US-0276484.
XX
PA   (CORI-) CORIXA CORP.
PA   (GAIG/) GAIGER A.
XX
PI   Gaiger A,  Cheever M;
XX
DR   WPI; 2000-293107/25.
XX
PT   Novel polypeptides comprising an immunogenic portion of a native WT1
PT   polypeptide, useful for inhibiting the development of malignant
PT   diseases associated with WT1 expression e.g. leukemia or cancer  -
XX
PS   Claim 4; Page 181; 193pp; English.
XX
CC   The present invention describes polypeptides (I) comprising an
CC   immunogenic portion of a native Wilm's Tumour gene product polypeptide,
CC   WT1, (or variants of the immunogenic portion retaining the ability to
CC   react with WT1-specific antisera and/or T-cell lines or clones) and
CC   comprising 16 consecutive amino acids (aa) or less of a native WT1
CC   polypeptide. The polypeptides are useful therapeutically and to
CC   manufacture medicaments for enhancing/inducing an immune response in
CC   patients. The polypeptides, mimetics or polynucleotides can be included
CC   with a carrier/excipient in pharmaceutical compositions or with a
CC   non-specific immune response enhancer (e.g. an adjuvant or enhancer
CC   preferentially enhancing a T cell response) in vaccines. Pharmaceutical
CC   compositions and vaccines can be administered to human patients to
CC   enhance or induce an immune response specific for WT1 or a cell
CC   expressing WT1, useful to inhibit the development of malignant diseases
CC   associated with WT1 expression, e.g. leukemia (especially acute/chronic
```

```
                    GenCore version 4.5
             Copyright (c) 1993 - 2000  Compugen Ltd.


OM protein - protein search, using sw model

Run on:         October 23, 2001, 13:21:01 ; Search time 46.09 Seconds
                                             (without alignments)
                                             11.838 Million cell updates/sec

Title:          US-09-625-963-2
Perfect score:  56
Sequence:       1 CMTWNQMNL 9

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched:       412676 seqs, 60623988 residues

Total number of hits satisfying chosen parameters:      412676

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :      A_Geneseq_0601:*
               1:  /SIDS8/gcgdata/geneseq/geneseqp/AA1980.DAT:*
               2:  /SIDS8/gcgdata/geneseq/geneseqp/AA1981.DAT:*
               3:  /SIDS8/gcgdata/geneseq/geneseqp/AA1982.DAT:*
               4:  /SIDS8/gcgdata/geneseq/geneseqp/AA1983.DAT:*
               5:  /SIDS8/gcgdata/geneseq/geneseqp/AA1984.DAT:*
               6:  /SIDS8/gcgdata/geneseq/geneseqp/AA1985.DAT:*
               7:  /SIDS8/gcgdata/geneseq/geneseqp/AA1986.DAT:*
               8:  /SIDS8/gcgdata/geneseq/geneseqp/AA1987.DAT:*
               9:  /SIDS8/gcgdata/geneseq/geneseqp/AA1988.DAT:*
               10:  /SIDS8/gcgdata/geneseq/geneseqp/AA1989.DAT:*
               11:  /SIDS8/gcgdata/geneseq/geneseqp/AA1990.DAT:*
               12:  /SIDS8/gcgdata/geneseq/geneseqp/AA1991.DAT:*
               13:  /SIDS8/gcgdata/geneseq/geneseqp/AA1992.DAT:*
               14:  /SIDS8/gcgdata/geneseq/geneseqp/AA1993.DAT:*
               15:  /SIDS8/gcgdata/geneseq/geneseqp/AA1994.DAT:*
               16:  /SIDS8/gcgdata/geneseq/geneseqp/AA1995.DAT:*
               17:  /SIDS8/gcgdata/geneseq/geneseqp/AA1996.DAT:*
               18:  /SIDS8/gcgdata/geneseq/geneseqp/AA1997.DAT:*
               19:  /SIDS8/gcgdata/geneseq/geneseqp/AA1998.DAT:*
               20:  /SIDS8/gcgdata/geneseq/geneseqp/AA1999.DAT:*
               21:  /SIDS8/gcgdata/geneseq/geneseqp/AA2000.DAT:*
               22:  /SIDS8/gcgdata/geneseq/geneseqp/AA2001.DAT:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                                 SUMMARIES
                  %
Result            Query
  No.    Score   Match Length DB  ID                           Description
 -----------------------------------------------------------------------------
    1       56   100.0      9  21  AAY94203                     Human cytotoxic T
    2       56   100.0      9  21  AAY98534                     WT1 derived immuno
    3       56   100.0      9  21  AAY98743                     WT1 derived immuno
    4       56   100.0      9  21  AAY80202                     Human Wilms' tumou
    5       56   100.0    345  12  AAR12240                     Wilm's tumour gene
    6       56   100.0    345  19  AAW47173                     Wilms' tumour poly
    7       56   100.0    429  16  AAR85065                     Wilms' tumour WT1
    8       56   100.0    449  19  AAW47176                     Wilms' tumour poly
    9       56   100.0    449  21  AAY98804                     Human WT1 protein
   10       56   100.0    449  21  AAY98805                     Mouse WT1 protein
   11       56   100.0    449  21  AAY80196                     Mouse Wilms' tumou
   12       56   100.0    449  21  AAY80197                     Human Wilms' tumou
   13       46    82.1      9  21  AAY98616                     WT1 derived immuno
   14       46    82.1      9  21  AAY98763                     WT1 derived immuno
   15       41    73.2      9  21  AAY98652                     WT1 derived immuno
   16       41    73.2    154  19  AAW47175                     Wilms' tumour poly
   17       37    66.1      9  21  AAY98730                     WT1 derived immuno
   18       37    66.1      9  21  AAY98790                     WT1 derived immuno
   19       34    60.7    262  14  AAR42502                     A. rhizogenes NIAE
   20       34    60.7    279  14  AAR42506                     Protein encoded by
   21       34    60.7    385  21  AAB19189                     Lipid metabolism p
   22       34    60.7    501  20  AAW93490                     Rat 1-alpha-OHase
   23       34    60.7    837  21  AAB12381                     N-terminal domain
   24       33    58.9     36  21  AAB52030                     Human secreted pro
   25       33    58.9     88  21  AAB08409                     Amino acid sequenc
   26       33    58.9    112  21  AAY65317                     Human 5' EST relat
   27       33    58.9    131  19  AAW42080                     The amino acid seq
   28       33    58.9    158  21  AAG48033                     Arabidopsis thalia
   29       33    58.9    171  21  AAG48032                     Arabidopsis thalia
   30       33    58.9    204  21  AAG32465                     Arabidopsis thalia
   31       33    58.9    204  21  AAG52115                     Arabidopsis thalia
   32       33    58.9    269  21  AAG18644                     Arabidopsis thalia
   33       33    58.9    277  21  AAG52114                     Arabidopsis thalia
   34       33    58.9    282  21  AAG18643                     Arabidopsis thalia
   35       33    58.9    282  21  AAG32464                     Arabidopsis thalia
   36       33    58.9    291  18  AAW35233                     Human tissue plasm
   37       33    58.9    298  21  AAG18642                     Arabidopsis thalia
   38       33    58.9    310   9  AAP80770                    Recombinant human
   39       33    58.9    314  10  AAP90174                     Tissue plasminogen
   40       33    58.9    337  22  AAB47084                     Adhesin protein, P
   41       33    58.9    345  10  AAP94415                     Sequence of coding
   42       33    58.9    345  10  AAP94414                     Sequence of coding
   43       33    58.9    348  14  AAR35401                     tPA-2.  Synthetic.
   44       33    58.9    348  14  AAR35404                     tPA-8.  Synthetic.
   45       33    58.9    349  11  AAR05511                     Plasminogen activa


                               ALIGNMENTS


RESULT   1
AAY94203
ID   AAY94203 standard; peptide; 9 AA.
XX
AC   AAY94203;
XX
DT   28-JUL-2000  (first entry)
XX
DE   Human cytotoxic T lymphocyte-recognised WT1 peptide WT235-43.
XX
KW   WT235-43; peptide; epitope; Wilm's tumour gene; leukaemia;
KW   breast cancer; melanoma; ovarian cancer; immunotherapy.
XX
OS   Homo sapiens.
XX
PN   WO200026249-A1.
XX
PD   11-MAY-2000.
XX
PF   02-NOV-1999;   99WO-GB03572.
XX
PR   02-NOV-1998;   98GB-0023897.
XX
PA   (IMCO-) IMPERIAL COLLEGE INNOVATIONS LTD.
XX
PI   Stauss HJ,  Gao L;
XX
DR   WPI; 2000-376123/32.
XX
PT   Novel peptides comprising WT-1 and GATA-1 epitopes, their fragments or
PT   variants, useful as vaccines for cancer immunotherapy  -
XX
PS   Claim 2; Page 74; 93pp; English.
XX
```

```
DR   Pfam; PF01485; IBR; 1.
DR   ProDom; PD000531; -; 1.
DR   PROSITE; PS00690; DEAH_ATP_HELICASE; 1.
DR   PROSITE; PS00290; IG_MHC; UNKNOWN_1.
DR   PROSITE; PS00028; ZINC_FINGER_C2H2; UNKNOWN_1.
DR   PROSITE; PS00518; ZINC_FINGER_C3HC4; UNKNOWN_2.
DR   SMART; SM00487; DEXDc; 1.
KW   Hypothetical protein.
SQ   SEQUENCE   2322 AA;  262590 MW;  D8BE0B7947108288 CRC64;

  Query Match             64.3%;  Score 36;  DB 10;  Length 2322;
  Best Local Similarity   62.5%;  Pred. No. 2.2e+02;
  Matches    5;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 CMTWNQMN 8
           | |||| :
Db      44 CTTWNQQH 51


RESULT  15
Q9SIA0
ID   Q9SIA0      PRELIMINARY;      PRT;   480 AA.
AC   Q9SIA0;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   AT2G04100 PROTEIN.
GN   AT2G04100.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC   Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. COLUMBIA;
RX   MEDLINE=20083487; PubMed=10617197;
RA   Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
RA   Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblyum T.V.,
RA   Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,
RA   Cronin L.A., Shen M., VanAken S.E., Umayam L., Tallon L.J., Gill J.E.,
RA   Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,
RA   Copenhaver G.P., Preuss D., Nierman W.C., White O., Eisen J.A.,
RA   Salzberg S.L., Fraser C.M., Venter J.C.;
RT   "Sequence and analysis of chromosome II of Arabidopsis thaliana.";
RL   Nature 402:761-768(1999).
DR   EMBL; AC007178; AAD28682.1; -.
DR   InterPro; IPR002528; -.
DR   Pfam; PF01554; UPF0013; 2.
SQ   SEQUENCE   480 AA;  51820 MW;  BE1B4C67660747FF CRC64;

  Query Match             63.4%;  Score 35.5;  DB 10;  Length 480;
  Best Local Similarity   60.0%;  Pred. No. 62;
  Matches    6;  Conservative    3;  Mismatches    0;  Indels    1;  Gaps    1;

Qy       1 CMTWN-QMNL 9
           ||:|| |:|:
Db     466 CMSWNEQVNI 475


Search completed: October 23, 2001, 13:27:54
Job time: 407 sec
```

```
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
DR   EMBL; AE003777; AAF57124.1; -.
DR   FlyBase; FBgn0039826; CG15560.
SQ   SEQUENCE   619 AA;  69911 MW;  08C3D5128BE30D13 CRC64;


  Query Match             66.1%;  Score 37;  DB 5;  Length 619;
  Best Local Similarity   75.0%;  Pred. No. 41;
  Matches    6;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      2 MTWNQMNL 9
          :|||| ||
Db    533 ITWNQYNL 540


RESULT  12
O62160
ID   O62160      PRELIMINARY;    PRT;   366 AA.
AC   O62160;
DT   01-AUG-1998 (TrEMBLrel. 07, Created)
DT   01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT   01-JAN-1999 (TrEMBLrel. 09, Last annotation update)
DE   F14B6.3 PROTEIN.
GN   F14B6.3.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   White S.;
RL   Submitted (NOV-1996) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94150718; PubMed=7906398;
RA   Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
RA   Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA   Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
RA   Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
RA   Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
RA   Lightning J., Lloyd C., Mcmurray A., Mortimore B., O'Callaghan M.,
RA   Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Shownkeen R.,
RA   Smaldon N., Smith A., Sonnhammer E., Staden R., Sulston J.,
RA   Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
RA   Watson A., Weinstock L., Wilkinson-Sproat J., Wohldman P.;
RT   "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT   elegans.";
RL   Nature 368:32-38(1994).
DR   EMBL; Z81502; CAB04104.1; -.
SQ   SEQUENCE   366 AA;  42102 MW;  C26A9C4AD944CC82 CRC64;


  Query Match             64.3%;  Score 36;  DB 5;  Length 366;
  Best Local Similarity   71.4%;  Pred. No. 38;
  Matches    5;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 CMTWNQM 7
          || ||:|
Db     74 CMRWNRM 80


RESULT  13
Q98383
ID   Q98383      PRELIMINARY;      PRT;   407 AA.
AC   Q98383;
DT   01-FEB-1997 (TrEMBLrel. 02, Created)
DT   01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   HEMAGGLUTININ (FRAGMENT).
GN   H.
OS   Measles virus.
OC   Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC   Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX   NCBI_TaxID=11234;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=X932;
RA   Outlaw M.C., Pringle C.R.;
RL   Submitted (JUL-1996) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U72221; AAB09682.1; -.
DR   InterPro; IPR000665; -.
DR   Pfam; PF00423; HN; 1.
FT   NON_TER       1      1
FT   NON_TER     407    407
SQ   SEQUENCE   407 AA;  45039 MW;  4E25F237B0D4BD91 CRC64;


  Query Match             64.3%;  Score 36;  DB 14;  Length 407;
  Best Local Similarity   83.3%;  Pred. No. 43;
  Matches    5;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 CMTWNQ 6
          | ||||
Db    360 CFTWNQ 365


RESULT  14
Q9SV27
ID   Q9SV27      PRELIMINARY;      PRT;  2322 AA.
AC   Q9SV27;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   HYPOTHETICAL 262.6 KDA PROTEIN.
GN   F3I3.40 OR AT4G01020.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC   Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Bevan M., Van Der Schueren J., Chuang Y.J., Voet M., Robben J.,
RA   Volckaert G., Bancroft I., Mewes H.W., Mayer K.F.X., Lemcke K.,
RA   Mannhaupt G., Schueller C.;
RL   Submitted (JUN-1999) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   EU Arabidopsis sequencing project;
RL   Submitted (JUN-1999) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Robben J., Grymonprez B., Weltjens I., Volckaert G., Mewes H.W.,
RA   Lemcke K., Mayer K.F.X.;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RA   EU Arabidopsis sequencing project;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AL080237; CAB45785.1; -.
DR   EMBL; AL161491; CAB80911.1; -.
DR   InterPro; IPR000822; -.
DR   InterPro; IPR000916; -.
DR   InterPro; IPR001410; -.
DR   InterPro; IPR001650; -.
DR   InterPro; IPR001841; -.
DR   InterPro; IPR002464; -.
DR   InterPro; IPR002867; -.
DR   InterPro; IPR003006; -.
DR   Pfam; PF00097; zf-C3HC4; 2.
DR   Pfam; PF00271; helicase_C; 1.
```

```
RT   -dramatic compaction and conservation of synteny with human chromosome
RT   11p13.";
RL   Submitted (JAN-1998) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.
DR   EMBL; AL021531; CAA16491.1; -.
DR   HSSP; P08046; 1AAY.
DR   InterPro; IPR000822; -.
DR   InterPro; IPR000976; -.
DR   Pfam; PF00096; zf-C2H2; 4.
DR   Pfam; PF02165; WT1; 2.
DR   PRINTS; PR00048; ZINCFINGER.
DR   PROSITE; PS00028; ZINC_FINGER_C2H2; 4.
DR   SMART; SM00355; ZnF_C2H2; 1.
KW   DNA-binding; Metal-binding; Nuclear protein; Zinc-finger.
SQ   SEQUENCE   416 AA;  45806 MW;  9C362E1684E67668 CRC64;

  Query Match             85.7%;  Score 48;  DB 13;  Length 416;
  Best Local Similarity   87.5%;  Pred. No. 0.25;
  Matches    7;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CMTWNQMN 8
           |:||||||
Db     216 CVTWNQMN 223


RESULT   9
Q9KAL6
ID   Q9KAL6      PRELIMINARY;      PRT;   310 AA.
AC   Q9KAL6;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   DIPEPTIDASE.
GN   BH2271.
OS   Bacillus halodurans.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group;
OC   Bacillus/Staphylococcus group; Bacillus.
OX   NCBI_TaxID=86665;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C-125 / JCM 9153;
RA   Takami H., Nakasone K., Takaki Y.;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AP001515; BAB05990.1; -.
DR   InterPro; IPR000180; -.
DR   PROSITE; PS00869; RENAL_DIPEPTIDASE; 1.
SQ   SEQUENCE   310 AA;  34778 MW;  28696A5A0FC89C0A CRC64;

  Query Match             67.9%;  Score 38;  DB 2;  Length 310;
  Best Local Similarity   75.0%;  Pred. No. 14;
  Matches    6;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 MTWNQMNL 9
             :|||| ||
Db     129 LTWNQANL 136


RESULT  10
Q9HLG8
ID   Q9HLG8      PRELIMINARY;      PRT;   285 AA.
AC   Q9HLG8;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   PROBABLE PYRUVATE FERREDOXIN OXIDOREDUCTASE, BETA SUBUNIT.
GN   TA0260.
OS   Thermoplasma acidophilum.
OC   Archaea; Euryarchaeota; Thermoplasmales; Thermoplasmaceae;
OC   Thermoplasma.
OX   NCBI_TaxID=2303;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=DSM 1728;
RX   MEDLINE=20479972; PubMed=11029001;
RA   Ruepp A., Graml W., Santos-Martinez M.-L., Koretke K.K., Volker C.,
RA   Mewes H.-W., Frishman D., Stocker S., Lupas A.N., Baumeister W.;
RT   "The genome sequence of the thermoacidophilic scavenger Thermoplasma
RT   acidophilum.";
RL   Nature 407:508-513(2000).
DR   EMBL; AL445063; CAC11405.1; -.
KW   Pyruvate.
SQ   SEQUENCE   285 AA;  31717 MW;  64F495E809C82F7A CRC64;

  Query Match             66.1%;  Score 37;  DB 1;  Length 285;
  Best Local Similarity   62.5%;  Pred. No. 20;
  Matches    5;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 CMTWNQMN 8
           |:|||: |
Db     202 CVTWNKEN 209


RESULT  11
Q9VA04
ID   Q9VA04      PRELIMINARY;      PRT;   619 AA.
AC   Q9VA04;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   CG15560 PROTEIN.
GN   CG15560.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BERKELEY;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
```

```
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   WT1(+KTS) PROTEIN.
GN   WT1(+KTS).
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=OVARY;
RA   Kudo T.;
RT   "Chicken counterpart of Wilms' tumor suppressor gene 1.";
RL   Submitted (OCT-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB033634; BAA94794.1; -.
DR   InterPro; IPR000822; -.
DR   InterPro; IPR000976; -.
DR   Pfam; PF00096; zf-C2H2; 4.
DR   Pfam; PF02165; WT1; 2.
DR   PRINTS; PR00049; WILMSTUMOUR.
DR   PROSITE; PS00028; ZINC_FINGER_C2H2; 4.
KW   DNA-binding; Metal-binding; Zinc-finger.
SQ   SEQUENCE   417 AA;  46957 MW;  47BBF7F6448E7F6C CRC64;

  Query Match             100.0%;  Score 56;  DB 13;  Length 417;
  Best Local Similarity   100.0%;  Pred. No. 0.0082;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CMTWNQMNL 9
           |||||||||
Db     218 CMTWNQMNL 226


RESULT   6
Q9W611
ID   Q9W611      PRELIMINARY;      PRT;   426 AA.
AC   Q9W611;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   WT1.
OS   Cynops pyrrhogaster (Japanese common newt).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Caudata; Salamandroidea; Salamandridae; Cynops.
OX   NCBI_TaxID=8330;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=TESTIS;
RA   Nakayama Y., Yamamoto T., Matsuda Y., Abe S.I.;
RT   "Cloning of cDNA for newt WT1 and the differential expression during
RT   spermatogenesis of the Japanese newt, Cynops pyrrhogaster.";
RL   Submitted (MAY-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB013888; BAA76399.1; -.
DR   HSSP; P08046; 1AAY.
DR   InterPro; IPR000822; -.
DR   InterPro; IPR000976; -.
DR   Pfam; PF00096; zf-C2H2; 4.
DR   Pfam; PF02165; WT1; 2.
DR   PRINTS; PR00049; WILMSTUMOUR.
DR   PROSITE; PS00028; ZINC_FINGER_C2H2; 4.
DR   SMART; SM00355; ZnF_C2H2; 1.
KW   DNA-binding; Metal-binding; Zinc-finger.
SQ   SEQUENCE   426 AA;  47590 MW;  46A0B158A9F79C6C CRC64;

  Query Match             100.0%;  Score 56;  DB 13;  Length 426;
  Best Local Similarity   100.0%;  Pred. No. 0.0083;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CMTWNQMNL 9
           |||||||||
Db     225 CMTWNQMNL 233
```

```
RESULT   7
Q15881
ID   Q15881      PRELIMINARY;      PRT;   446 AA.
AC   Q15881;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   WILMS TUMOR PROTEIN.
GN   WT1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92241883; PubMed=1572653;
RA   Gessler M., Konig A., Bruns G.A.;
RT   "The genomic organization and expression of the WT1 gene.";
RL   Genomics 12:807-813(1992).
CC   -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.
DR   EMBL; X61631; CAA43819.1; -.
DR   EMBL; X61632; CAA43819.1; JOINED.
DR   EMBL; X61633; CAA43819.1; JOINED.
DR   EMBL; X61634; CAA43819.1; JOINED.
DR   EMBL; X61635; CAA43819.1; JOINED.
DR   EMBL; X61636; CAA43819.1; JOINED.
DR   EMBL; X61637; CAA43819.1; JOINED.
DR   EMBL; X61638; CAA43819.1; JOINED.
DR   HSSP; P08046; 1AAY.
DR   InterPro; IPR000822; -.
DR   InterPro; IPR000976; -.
DR   Pfam; PF00096; zf-C2H2; 4.
DR   Pfam; PF02165; WT1; 1.
DR   PRINTS; PR00049; WILMSTUMOUR.
DR   PRINTS; PR00048; ZINCFINGER.
DR   PROSITE; PS00028; ZINC_FINGER_C2H2; 4.
DR   SMART; SM00355; ZnF_C2H2; 1.
KW   DNA-binding; Metal-binding; Nuclear protein; Zinc-finger.
SQ   SEQUENCE   446 AA;  48842 MW;  8CE7FC047F41CF11 CRC64;

  Query Match             100.0%;  Score 56;  DB 4;  Length 446;
  Best Local Similarity   100.0%;  Pred. No. 0.0087;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CMTWNQMNL 9
           |||||||||
Db     235 CMTWNQMNL 243


RESULT   8
O93433
ID   O93433      PRELIMINARY;      PRT;   416 AA.
AC   O93433;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   WILMS TUMOUR GENE.
GN   WT1.
OS   Fugu rubripes (Japanese pufferfish) (Takifugu rubripes).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC   Tetraodontidae; Takifugu.
OX   NCBI_TaxID=31033;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Miles C., Elgar G., Coles E., Kleinjan D.J., Vanheyningen V.,
RA   Hastie N.;
RT   "Complete sequencing of the Fugu WAGR region from WT1 to PAX6
```

```
Db    218 CMTWNQMNL 226


RESULT   2
P79958
ID   P79958      PRELIMINARY;      PRT;   407 AA.
AC   P79958;
DT   01-MAY-1997 (TrEMBLrel. 03, Created)
DT   01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   WT1 PROTEIN.
GN   WT1.
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=TESTIS;
RX   MEDLINE=97074667; PubMed=8917094;
RA   Semba K., Saito-Ueno R., Takayama G., Kondo M.;
RT   "cDNA cloning and its pronephros-specific expression of the Wilms'
RT   tumor suppressor gene, WT1, from Xenopus laevis.";
RL   Gene 175:167-172(1996).
CC   -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.
DR   EMBL; D82051; BAA11522.1; -.
DR   HSSP; P08046; 1AAY.
DR   InterPro; IPR000822; -.
DR   InterPro; IPR000976; -.
DR   Pfam; PF00096; zf-C2H2; 4.
DR   Pfam; PF02165; WT1; 3.
DR   PRINTS; PR00048; ZINCFINGER.
DR   PROSITE; PS00028; ZINC_FINGER_C2H2; 4.
DR   SMART; SM00355; ZnF_C2H2; 1.
KW   DNA-binding; Metal-binding; Nuclear protein; Zinc-finger.
SQ   SEQUENCE   407 AA;  45983 MW;  E2554C658005870C CRC64;


  Query Match             100.0%;  Score 56;  DB 13;  Length 407;
  Best Local Similarity   100.0%;  Pred. No. 0.008;
  Matches    9; Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 CMTWNQMNL 9
          |||||||||
Db    212 CMTWNQMNL 220


RESULT   3
Q91657
ID   Q91657      PRELIMINARY;      PRT;   409 AA.
AC   Q91657;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   WILMS' TUMOR SUPPRESSOR (WT1).
GN   WT1.
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Carroll T.J., Vize P.D.;
RL   Submitted (JAN-1996) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE OF 289-370 FROM N.A.
RC   TISSUE=MESONEPHROS;
RX   MEDLINE=96068905; PubMed=7478606;
RA   Kent J., Coriat A.M., Sharpe P.T., Hastie N., Van Heyningen V.;
RT   "The evolution of WT1 sequence and expression pattern in the
RT   vertebrates.";
RL   Oncogene 11:1781-1792(1995).
CC   -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.
DR   EMBL; U42011; AAB53152.1; -.
DR   EMBL; X85733; CAA59738.1; -.
DR   HSSP; P08046; 1AAY.
DR   InterPro; IPR000822; -.
DR   InterPro; IPR000976; -.
DR   Pfam; PF00096; zf-C2H2; 4.
DR   Pfam; PF02165; WT1; 3.
DR   PRINTS; PR00048; ZINCFINGER.
DR   PROSITE; PS00028; ZINC_FINGER_C2H2; 4.
DR   SMART; SM00355; ZnF_C2H2; 1.
KW   DNA-binding; Metal-binding; Nuclear protein; Zinc-finger.
SQ   SEQUENCE   409 AA;  46186 MW;  2217FC04612CDE10 CRC64;

  Query Match             100.0%;  Score 56;  DB 13;  Length 409;
  Best Local Similarity   100.0%;  Pred. No. 0.008;
  Matches    9; Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 CMTWNQMNL 9
          |||||||||
Db    214 CMTWNQMNL 222


RESULT   4
Q9I8A1
ID   Q9I8A1      PRELIMINARY;      PRT;   414 AA.
AC   Q9I8A1;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   WT1(-KTS) PROTEIN.
GN   WT1(-KTS).
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=OVARY;
RA   Kudo T.;
RT   "Chicken counterpart of Wilms' tumor suppressor gene 1.";
RL   Submitted (OCT-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB033633; BAA94793.1; -.
DR   InterPro; IPR000822; -.
DR   InterPro; IPR000976; -.
DR   Pfam; PF00096; zf-C2H2; 4.
DR   Pfam; PF02165; WT1; 2.
DR   PRINTS; PR00049; WILMSTUMOUR.
DR   PROSITE; PS00028; ZINC_FINGER_C2H2; 4.
KW   DNA-binding; Metal-binding; Zinc-finger.
SQ   SEQUENCE   414 AA;  46641 MW;  17EB22F8B428A2EF CRC64;


  Query Match             100.0%;  Score 56;  DB 13;  Length 414;
  Best Local Similarity   100.0%;  Pred. No. 0.0081;
  Matches    9; Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 CMTWNQMNL 9
          |||||||||
Db    218 CMTWNQMNL 226

RESULT   5
Q9I8A0
ID   Q9I8A0      PRELIMINARY;      PRT;   417 AA.
AC   Q9I8A0;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
```

```
                        GenCore version 4.5
              Copyright (c) 1993 - 2000  Compugen Ltd.

OM protein - protein search, using sw model

Run on:          October 23, 2001, 13:27:54 ; Search time 47.11 Seconds
                                              (without alignments)
                                              25.276 Million cell updates/sec

Title:           US-09-625-963-2
Perfect score:   56
Sequence:        1 CMTWNQMNL 9

Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched:        425026 seqs, 132305027 residues

Total number of hits satisfying chosen parameters:      425026

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       SPTREMBL_16:*
                1:  sp_archea:*
                2:  sp_bacteria:*
                3:  sp_fungi:*
                4:  sp_human:*
                5:  sp_invertebrate:*
                6:  sp_mammal:*
                7:  sp_mhc:*
                8:  sp_organelle:*
                9:  sp_phage:*
                10:  sp_plant:*
                11:  sp_rodent:*
                12:  sp_unclassified:*
                13:  sp_vertebrate:*
                14:  sp_virus:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                              SUMMARIES
                 %
Result           Query
  No.    Score   Match Length DB  ID                            Description
 -------------------------------------------------------------------------------
     1      56   100.0    392  13  Q91030                        Q91030 gallus gall
     2      56   100.0    407  13  P79958                        P79958 xenopus lae
     3      56   100.0    409  13  Q91657                        Q91657 xenopus lae
     4      56   100.0    414  13  Q9I8A1                        Q9i8a1 gallus gall
     5      56   100.0    417  13  Q9I8A0                        Q9i8a0 gallus gall
     6      56   100.0    426  13  Q9W611                        Q9w611 cynops pyrr
     7      56   100.0    446   4  Q15881                       Q15881 homo sapien
     8      48    85.7    416  13  O93433                        O93433 fugu rubrip
     9      38    67.9    310   2  Q9KAL6                       Q9kal6 bacillus ha
    10      37    66.1    285   1  Q9HLG8                       Q9hlg8 thermoplasm
    11      37    66.1    619   5  Q9VA04                       Q9va04 drosophila
    12      36    64.3    366   5  O62160                       O62160 caenorhabdi
    13      36    64.3    407  14  Q98383                        Q98383 measles vir
    14      36    64.3   2322  10  Q9SV27                        Q9sv27 arabidopsis
    15    35.5    63.4    480  10  Q9SIA0                        Q9sia0 arabidopsis
    16      35    62.5    183  10  Q9LFV0                        Q9lfv0 arabidopsis
    17      35    62.5    186   8  Q34247                       Q34247 ceratovacun
    18      35    62.5    345   8  Q9TG93                       Q9tg93 anniella pu
    19      35    62.5    347   5  Q9NAN5                       Q9nan5 caenorhabdi
    20      35    62.5    369   9  Q9XJ97                       Q9xj97 streptococc
    21      35    62.5    419  13  Q9PUT7                        Q9put7 brachydanio
    22      35    62.5    544   5  Q9VTV0                       Q9vtv0 drosophila
    23      35    62.5    545   5  O16180                       O16180 drosophila
    24      35    62.5    545   5  O16169                       O16169 drosophila
    25      35    62.5    545   5  O16171                       O16171 drosophila
    26      35    62.5    548   5  O16177                       O16177 drosophila
    27      35    62.5    548   5  O16173                       O16173 drosophila
    28      35    62.5    555   5  O16168                       O16168 drosophila
    29      35    62.5    623   9  Q9XJW0                       Q9xjw0 streptococc
    30      35    62.5    623   9  Q9XJ75                       Q9xj75 streptococc
    31      35    62.5    661   5  Q9W435                       Q9w435 drosophila
    32      35    62.5    694  10  Q9SVK2                        Q9svk2 arabidopsis
    33      35    62.5    747   5  Q9TZL0                       Q9tzl0 caenorhabdi
    34      35    62.5    849  10  O81906                        O81906 arabidopsis
    35      35    62.5    850  10  O04384                        O04384 brassica ol
    36      35    62.5   1124   5  O45611                       O45611 caenorhabdi
    37      35    62.5   1142   5  Q9GNN7                       Q9gnn7 caenorhabdi
    38      35    62.5   1368   5  Q23821                       Q23821 caenorhabdi
    39      35    62.5   1749  11  Q9JKU3                        Q9jku3 rattus norv
    40      35    62.5   2294   5  Q9N3X8                       Q9n3x8 caenorhabdi
    41    34.5    61.6    580   5  Q9V8L4                       Q9v814 drosophila
    42      34    60.7     79   2  Q9KJ55                       Q9kj55 lactococcus
    43      34    60.7     79   9  Q9MBY5                       Q9mby5 lactococcus
    44      34    60.7    141  11  O70429                        O70429 rattus norv
    45      34    60.7    156   2  Q9K3K1                       Q9k3k1 streptomyce

                              ALIGNMENTS


RESULT  1
Q91030
ID   Q91030      PRELIMINARY;      PRT;   392 AA.
AC   Q91030;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   CHICK WILM'S TUMOUR PROTEIN (FRAGMENT).
GN   CWT1.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=WHOLE EMBRYO;
RX   MEDLINE=96068905; PubMed=7478606;
RA   Kent J., Coriat A.M., Sharpe P.T., Hastie N., Van Heyningen V.;
RT   "The evolution of WT1 sequence and expression pattern in the
RT   vertebrates.";
RL   Oncogene 11:1781-1792(1995).
DR   EMBL; X85731; CAA59736.1; -.
DR   HSSP; P08046; 1A1G.
DR   InterPro; IPR000822; -.
DR   InterPro; IPR000976; -.
DR   Pfam; PF00096; zf-C2H2; 3.
DR   Pfam; PF02165; WT1; 2.
DR   PRINTS; PR00049; WILMSTUMOUR.
DR   PROSITE; PS00028; ZINC_FINGER_C2H2; 3.
DR   SMART; SM00355; ZnF_C2H2; 1.
KW   DNA-binding; Metal-binding; Zinc-finger.
FT   NON_TER     392    392
SQ   SEQUENCE   392 AA;  43869 MW;  761F01D350E4EDBF CRC64;


  Query Match             100.0%;  Score 56;  DB 13;  Length 392;
  Best Local Similarity   100.0%;  Pred. No. 0.0077;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CMTWNQMNL 9
           |||||||||
```

```
Qy      1 CMTWNQM 7
          | ||||:
Db    605 CDTWNQL 611


RESULT  15
CDR3_CANAL
ID   CDR3_CANAL     STANDARD;      PRT;  1501 AA.
AC   O42690;
DT   15-JUL-1999 (Rel. 38, Created)
DT   15-JUL-1999 (Rel. 38, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   OPAQUE-SPECIFIC ABC TRANSPORTER CDR3.
GN   CDR3.
OS   Candida albicans (Yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; mitosporic Saccharomycetales; Candida.
OX   NCBI_TaxID=5476;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=1006;
RX   MEDLINE=98053831; PubMed=9393682;
RA   Balan I., Alarco A.M., Raymond M.;
RT   "The Candida albicans CDR3 gene codes for an opaque-phase ABC
RT   transporter.";
RL   J. Bacteriol. 179:7210-7218(1997).
CC   -!- FUNCTION: NOT KNOWN.
CC   -!- DEVELOPMENTAL STAGE: REGULATED IN A CELL-TYPE-SPECIFIC MANNER;
CC       HIGH LEVELS IN WO-1 OPAQUE CELLS, AND UNDETECTABLE LEVELS IN WO-1
CC       WHITE CELLS.
CC   -!- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY
CC       (ABC TRANSPORTERS). PDR5 SUBFAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; U89714; AAC49886.1; -.
DR   InterPro; IPR001617; -.
DR   Pfam; PF00005; ABC_tran; 2.
DR   PROSITE; PS00211; ABC_TRANSPORTER; 1.
KW   ATP-binding; Transmembrane; Glycoprotein; Transport; Repeat.
FT   DOMAIN        1    502       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    503    523       POTENTIAL.
FT   TRANSMEM    540    560       POTENTIAL.
FT   TRANSMEM    589    609       POTENTIAL.
FT   TRANSMEM    614    634       POTENTIAL.
FT   TRANSMEM    653    673       POTENTIAL.
FT   TRANSMEM    755    775       POTENTIAL.
FT   DOMAIN      776   1175       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM   1176   1196       POTENTIAL.
FT   TRANSMEM   1212   1232       POTENTIAL.
FT   TRANSMEM   1261   1281       POTENTIAL.
FT   TRANSMEM   1297   1317       POTENTIAL.
FT   TRANSMEM   1325   1345       POTENTIAL.
FT   TRANSMEM   1451   1471       POTENTIAL.
FT   NP_BIND     876    883       ATP (POTENTIAL).
SQ   SEQUENCE   1501 AA;  170270 MW;  08A5C43DA97C06DE CRC64;


  Query Match             62.5%;  Score 35;  DB 1;  Length 1501;
  Best Local Similarity   57.1%;  Pred. No. 1.4e+02;
  Matches    4;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 CMTWNQM 7
          |::|||:
Db   1316 CISWNQL 1322
```

```
Search completed: October 23, 2001, 13:28:18
Job time: 401 sec
```

```
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7237;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91133260; PubMed=2178209;
RA   Brady J.P., Richmond R.C., Oakeshott J.G.;
RT   "Cloning of the esterase-5 locus from Drosophila pseudoobscura and
RT   comparison with its homologue in D. melanogaster.";
RL   Mol. Biol. Evol. 7:525-546(1990).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JR198E, GB4E, JR50E, GB8E, GB115E, GB139E, AND GB336E;
RX   MEDLINE=98133888; PubMed=9475741;
RA   King L.M.;
RT   "The role of gene conversion in determining sequence variation and
RT   divergence in the Est-5 gene family in Drosophila pseudoobscura.";
RL   Genetics 148:305-315(1998).
CC   -!- CATALYTIC ACTIVITY: A CARBOXYLIC ESTER + H(2)O = AN ALCOHOL
CC       + A CARBOXYLIC ANION.
CC   -!- SIMILARITY: BELONGS TO THE TYPE-B CARBOXYLESTERASE/LIPASE FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; M55907; AAA28514.1; -.
DR   EMBL; AF016144; AAB70234.1; -.
DR   EMBL; AF016145; AAB70235.1; -.
DR   EMBL; AF016146; AAB70236.1; -.
DR   EMBL; AF016147; AAB70237.1; -.
DR   EMBL; AF016148; AAB70238.1; -.
DR   EMBL; AF016149; AAB70239.1; -.
DR   EMBL; AF016150; AAB70240.1; -.
DR   HSSP; P21836; 1MAH.
DR   FlyBase; FBgn0015146; Dpse\Est-5C.
DR   InterPro; IPR002018; -.
DR   Pfam; PF00135; COesterase; 1.
DR   PROSITE; PS00122; CARBOXYLESTERASE_B_1; 1.
DR   PROSITE; PS00941; CARBOXYLESTERASE_B_2; 1.
KW   Hydrolase; Serine esterase; Glycoprotein; Signal.
FT   SIGNAL        1     19       POTENTIAL.
FT   CHAIN        20    545       ESTERASE 5C.
FT   ACT_SITE    207    207       BY SIMILARITY.
FT   ACT_SITE    467    467       BY SIMILARITY.
FT   DISULFID     84    103       BY SIMILARITY.
FT   DISULFID    259    271       BY SIMILARITY.
FT   DISULFID    515    536       POTENTIAL.
FT   CARBOHYD    113    113       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    421    421       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    507    507       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARIANT       6      6       L -> R (IN STRAIN GB8E).
FT   VARIANT      11     11       S -> R (IN STRAIN GB4E).
FT   VARIANT      42     42       F -> N (IN STRAINS GB8E AND GB336E).
FT   VARIANT      58     58       E -> D (IN STRAINS JR50E AND GB139E).
FT   VARIANT      98     98       V -> A (IN STRAINS JR50E, GB8E, AND
FT                                GB115E).
FT   VARIANT     170    170       D -> G (IN STRAINS GB4E, JR50E, GB115E,
FT                                GB139E, AND GB336E).
FT   VARIANT     266    266       M -> T (IN STRAINS GB4E, JR50E, GB115E,
FT                                GB139E, AND GB336E).
FT   VARIANT     291    291       A -> G (IN STRAIN JR50E).
FT   VARIANT     307    307       A -> T (IN STRAIN GB115E).
FT   VARIANT     461    461       V -> I (IN STRAIN GB336E).
FT   VARIANT     537    537       E -> V (IN STRAIN JR50E).
SQ   SEQUENCE   545 AA;  61021 MW;  60535F4DFA639D6A CRC64;

  Query Match              62.5%;  Score 35;  DB 1;  Length 545;
```

```
  Best Local Similarity   83.3%;  Pred. No. 53;
  Matches    5;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 CMTWNQ 6
              || |||
Db         84 CMQWNQ 89


RESULT 14
PRTD_YEAST
ID   PRTD_YEAST     STANDARD;      PRT;   712 AA.
AC   P25375;
DT   01-MAY-1992 (Rel. 22, Created)
DT   01-MAY-1992 (Rel. 22, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   SACCHAROLYSIN (EC 3.4.24.37) (PROTEASE D) (PROTEINASE YSCD)
DE   (OLIGOPEPTIDASE YSCD).
GN   PRD1 OR YCL057W OR YCL57W.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94139744; PubMed=8307027;
RA   Buechler M., Tisljar U., Wolf D.H.;
RT   "Proteinase yscD (oligopeptidase yscD). Structure, function and
RT   relationship of the yeast enzyme with mammalian thimet oligopeptidase
RT   (metalloendopeptidase, EP 24.15).";
RL   Eur. J. Biochem. 219:627-639(1994).
RN   [2]
RP   SEQUENCE OF 1-663 FROM N.A.
RA   Alexandraki D., Galland P., Thireos G., Tzermia M.;
RL   Submitted (MAR-1992) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE OF 663-712 FROM N.A.
RA   Voet M., Volckaert G.;
RL   Submitted (MAR-1992) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: COULD BE INVOLVED IN LATE STAGE OF PROTEIN DEGRADATION.
CC   -!- CATALYTIC ACTIVITY: CLEAVAGE OF PRO-|-PHE AND ALA-|-ALA BONDS.
CC   -!- COFACTOR: BINDS ONE ZINC ION (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M3 (ZINC METALLOPROTEASE);
CC       ALSO KNOWN AS THE THIMET OLIGOPEPTIDASE SUBFAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; X76504; CAA54039.1; -.
DR   EMBL; X59720; CAA42388.1; -.
DR   PIR; S19387; S19387.
DR   MEROPS; M03.003; -.
DR   SGD; S0000562; PRD1.
DR   InterPro; IPR000130; -.
DR   InterPro; IPR001567; -.
DR   Pfam; PF01432; Peptidase_M3; 1.
DR   PROSITE; PS00142; ZINC_PROTEASE; 1.
KW   Hydrolase; Metalloprotease; Zinc.
FT   METAL       501    501       ZINC (CATALYTIC) (BY SIMILARITY).
FT   ACT_SITE    502    502       BY SIMILARITY.
FT   METAL       505    505       ZINC (CATALYTIC) (BY SIMILARITY).
FT   METAL       508    508       ZINC (CATALYTIC) (BY SIMILARITY).
SQ   SEQUENCE   712 AA;  81933 MW;  340910B7FDAFBE37 CRC64;

  Query Match              62.5%;  Score 35;  DB 1;  Length 712;
  Best Local Similarity   71.4%;  Pred. No. 68;
  Matches    5;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;
```

```
RESULT  11
ESTA_DROPS
ID   ESTA_DROPS     STANDARD;      PRT;   544 AA.
AC   P25727;
DT   01-MAY-1992 (Rel. 22, Created)
DT   01-MAY-1992 (Rel. 22, Last sequence update)
DT   01-OCT-1996 (Rel. 34, Last annotation update)
DE   ESTERASE 5A PRECURSOR (EC 3.1.1.1) (EST-5A) (CARBOXYLIC-ESTER
DE   HYDROLASE).
GN   EST5A.
OS   Drosophila pseudoobscura (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7237;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91133260; PubMed=2178209;
RA   Brady J.P., Richmond R.C., Oakeshott J.G.;
RT   "Cloning of the esterase-5 locus from Drosophila pseudoobscura and
RT   comparison with its homologue in D. melanogaster.";
RL   Mol. Biol. Evol. 7:525-546(1990).
CC   -!- CATALYTIC ACTIVITY: A CARBOXYLIC ESTER + H(2)O = AN ALCOHOL
CC       + A CARBOXYLIC ANION.
CC   -!- SIMILARITY: BELONGS TO THE TYPE-B CARBOXYLESTERASE/LIPASE FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; M55908; AAA28516.1; -.
DR   HSSP; P21836; 1MAH.
DR   FlyBase; FBgn0015144; Dpse\Est-5A.
DR   InterPro; IPR002018; -.
DR   Pfam; PF00135; COesterase; 1.
DR   PROSITE; PS00122; CARBOXYLESTERASE_B_1; 1.
DR   PROSITE; PS00941; CARBOXYLESTERASE_B_2; 1.
KW   Hydrolase; Serine esterase; Glycoprotein; Signal.
FT   SIGNAL        1     21       POTENTIAL.
FT   CHAIN        22    544       ESTERASE 5A.
FT   ACT_SITE    209    209       BY SIMILARITY.
FT   DISULFID     86    105       BY SIMILARITY.
FT   DISULFID    261    273       BY SIMILARITY.
FT   DISULFID    514    535       POTENTIAL.
FT   CARBOHYD     94     94       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    115    115       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    475    475       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   544 AA;  61051 MW;  34A2023B3F2C26DB CRC64;

  Query Match             62.5%;  Score 35;  DB 1;  Length 544;
  Best Local Similarity   83.3%;  Pred. No. 53;
  Matches    5;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 CMTWNQ 6
           || |||
Db      86 CMQWNQ 91


RESULT  12
ESTP_DROME
ID   ESTP_DROME     STANDARD;      PRT;   544 AA.
AC   P18167;
DT   01-NOV-1990 (Rel. 16, Created)
DT   01-NOV-1990 (Rel. 16, Last sequence update)
DT   01-FEB-1996 (Rel. 33, Last annotation update)
DE   ESTERASE P PRECURSOR (EC 3.1.1.1) (EST-P) (CARBOXYLIC-ESTER
DE   HYDROLASE).
GN   ESTP OR EST-P.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CANTON-S;
RX   MEDLINE=90136038; PubMed=2105433;
RA   Collet C., Nielsen K.M., Russell R.J., Karl M., Oakeshott J.G.,
RA   Richmond R.C.;
RT   "Molecular analysis of duplicated esterase genes in Drosophila
RT   melanogaster.";
RL   Mol. Biol. Evol. 7:9-28(1990).
CC   -!- CATALYTIC ACTIVITY: A CARBOXYLIC ESTER + H(2)O = AN ALCOHOL
CC       + A CARBOXYLIC ANION.
CC   -!- SUBUNIT: MONOMER.
CC   -!- DEVELOPMENTAL STAGE: MAINLY IN LATE LARVAE.
CC   -!- SIMILARITY: BELONGS TO THE TYPE-B CARBOXYLESTERASE/LIPASE FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; M33780; AAA28520.1; -.
DR   PIR; B34089; B34089.
DR   HSSP; P21836; 1MAH.
DR   FlyBase; FBgn0000594; Est-P.
DR   InterPro; IPR002018; -.
DR   Pfam; PF00135; COesterase; 1.
DR   PROSITE; PS00122; CARBOXYLESTERASE_B_1; 1.
DR   PROSITE; PS00941; CARBOXYLESTERASE_B_2; 1.
KW   Hydrolase; Serine esterase; Glycoprotein; Signal.
FT   SIGNAL        1     19
FT   CHAIN        20    544       ESTERASE P.
FT   ACT_SITE    206    206       BY SIMILARITY.
FT   ACT_SITE    466    466       BY SIMILARITY.
FT   DISULFID     83    102       BY SIMILARITY.
FT   DISULFID    258    270       BY SIMILARITY.
FT   DISULFID    514    535       POTENTIAL.
FT   CARBOHYD     75     75       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    114    114       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    262    262       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    456    456       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   544 AA;  61230 MW;  6B60C5D9468CA43D CRC64;

  Query Match             62.5%;  Score 35;  DB 1;  Length 544;
  Best Local Similarity   83.3%;  Pred. No. 53;
  Matches    5;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 CMTWNQ 6
           || |||
Db      83 CMQWNQ 88


RESULT  13
ESTC_DROPS
ID   ESTC_DROPS     STANDARD;      PRT;   545 AA.
AC   P25725; O16181; O16182; O16183; O16184; O16185; O16186;
DT   01-MAY-1992 (Rel. 22, Created)
DT   01-MAY-1992 (Rel. 22, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   ESTERASE 5C PRECURSOR (EC 3.1.1.1) (EST-5C) (CARBOXYLIC-ESTER
DE   HYDROLASE) (CARBOXYLESTERASE-5C).
GN   EST5C.
OS   Drosophila pseudoobscura (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
```

```
CC   use  by  non-profit  institutions as long  as its content  is  in  no way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------------
DR   EMBL; X71395; CAA50517.1; -.
DR   HSSP; P00404; 1OCC.
DR   InterPro; IPR001505; -.
DR   InterPro; IPR002429; -.
DR   Pfam; PF00116; COX2; 1.
DR   PRINTS; PR01166; CYCOXIDASEII.
DR   PROSITE; PS00078; COX2; 1.
KW   Oxidoreductase; Copper; Mitochondrion; Transmembrane; Inner membrane;
KW   Electron transport; Respiratory chain.
FT   DOMAIN        1     26       MITOCHONDRIAL INTERMEMBRANE (POTENTIAL).
FT   TRANSMEM     27     48       PROBABLE.
FT   DOMAIN       49     62       MITOCHONDRIAL MATRIX (POTENTIAL).
FT   TRANSMEM     63     82       PROBABLE.
FT   DOMAIN       83    224       MITOCHONDRIAL INTERMEMBRANE (POTENTIAL).
FT   METAL       161    161       COPPER A (PROBABLE).
FT   METAL       196    196       COPPER A (PROBABLE).
FT   METAL       200    200       COPPER A (PROBABLE).
FT   METAL       204    204       COPPER A (PROBABLE).
SQ   SEQUENCE   224 AA;  25091 MW;  5D44CC670932075A CRC64;

  Query Match              64.3%;  Score 36;  DB 1;  Length 224;
  Best Local Similarity    85.7%;  Pred. No. 16;
  Matches    6;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          3 TWNQMNL 9
              || ||||
Db          3 TWGQMNL 9


RESULT   9
YJF5_YEAST
ID   YJF5_YEAST     STANDARD;      PRT;   245 AA.
AC   P47044;
DT   01-FEB-1996 (Rel. 33, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   HYPOTHETICAL 26.9 KDA PROTEIN IN BTN1-PEP8 INTERGENIC REGION.
GN   YJL055W OR J1148.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Pohl T.M., Aljinovic G.;
RL   Submitted (SEP-1995) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: TO R.FASCIANS HYPOTHETICAL 21.1 KDA PROTEIN IN
CC       FASCIATION LOCUS (AC P46378).
CC   -----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------------
DR   EMBL; Z49330; CAA89346.1; -.
DR   SGD; S0003591; YJL055W.
KW   Hypothetical protein; ATP-binding.
FT   NP_BIND      13     20       ATP (POTENTIAL).
SQ   SEQUENCE   245 AA;  26864 MW;  0C49E7FC4F1E7CA2 CRC64;

  Query Match              64.3%;  Score 36;  DB 1;  Length 245;
  Best Local Similarity    44.4%;  Pred. No. 17;
  Matches    4;  Conservative    4;  Mismatches    1;  Indels    0;  Gaps    0;
```

```
Qy          1 CMTWNQMNL 9
              |:||:|: :
Db        160 CITWSQLGI 168


RESULT  10
SSK2_YEAST
ID   SSK2_YEAST     STANDARD;      PRT;  1579 AA.
AC   P53599;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   SERINE/THREONINE PROTEIN KINASE SSK2 (EC 2.7.-.-) (SUPPRESSOR OF
DE   SENSOR KINASE 2).
GN   SSK2 OR YNR031C OR N3276.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=95350642; PubMed=7624781;
RA   Maeda T., Takekawa M., Saito H.;
RT   "Activation of yeast PBS2 MAPKK by MAPKKKs or by binding of an SH3-
RT   containing osmosensor.";
RL   Science 269:554-558(1995).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Pohl T.M.;
RL   Submitted (MAY-1996) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: KINASE INVOLVED IN A SIGNAL TRANSDUCTION PATHWAY THAT IS
CC       ACTIVATED BY CHANGES IN THE OSMOLARITY OF THE EXTRACELLULAR
CC       ENVIRONMENT. ACTIVATES THE PBS2 MAP KINASE KINASE BY
CC       PHOSPHORYLATION.
CC   -!- SUBUNIT: INTERACTS WITH BY SSK1.
CC   -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC       MAP KINASE KINASE KINASE SUBFAMILY.
CC   -----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------------
DR   EMBL; L41927; AAC41665.1; -.
DR   EMBL; Z71646; CAA96311.1; -.
DR   HSSP; P24941; 1AQ1.
DR   SGD; S0005314; SSK2.
DR   InterPro; IPR000719; -.
DR   InterPro; IPR002290; -.
DR   Pfam; PF00069; pkinase; 1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW   Transferase; Serine/threonine-protein kinase; ATP-binding.
FT   DOMAIN     1266   1558       PROTEIN KINASE.
FT   NP_BIND    1272   1280       ATP (BY SIMILARITY).
FT   BINDING    1295   1295       ATP (BY SIMILARITY).
FT   ACT_SITE   1390   1390       BY SIMILARITY.
SQ   SEQUENCE   1579 AA;  180526 MW;  91A07A195CB3772E CRC64;

  Query Match              64.3%;  Score 36;  DB 1;  Length 1579;
  Best Local Similarity    71.4%;  Pred. No. 97;
  Matches    5;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 TWNQMNL 9
              |||::||
Db        609 TWNELNL 615
```

```
RA   Tsurutani N., Oda H., Nakatsuru Y., Imai Y., Zhang S., Ueno Y.,
RA   Ishikawa T.;
RT   "cDNA cloning and developmental expression of the porcine homologue
RT   of WT1.";
RL   Gene 211:215-220(1998).
CC   -!- FUNCTION: POTENTIAL ROLE IN TRANSCRIPTIONAL REGULATION. RECOGNIZES
CC       AND BINDS TO THE DNA SEQUENCE 5'-CGCCCCCGC-3'.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR.
CC   -!- ALTERNATIVE PRODUCTS: TWO ALTERNATIVE SPLICE SITES EXISTS.
CC   -!- DEVELOPMENTAL STAGE: EXPRESSED DURING KIDNEY DEVELOPMENT.
CC   -!- SIMILARITY: BELONGS TO THE EGR FAMILY OF C2H2-TYPE ZINC-FINGER
CC       PROTEINS.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AB010969; BAA28147.1; -.
DR   InterPro; IPR000822; -.
DR   InterPro; IPR000976; -.
DR   Pfam; PF00096; zf-C2H2; 4.
DR   PRINTS; PR00048; ZINCFINGER.
DR   PRINTS; PR00049; WILMSTUMOUR.
DR   PROSITE; PS00028; ZINC_FINGER_C2H2_1; 4.
DR   PROSITE; PS50157; ZINC_FINGER_C2H2_2; 4.
KW   Zinc-finger; Metal-binding; DNA-binding; Repeat; Nuclear protein;
KW   Transcription regulation; Alternative splicing; Anti-oncogene.
FT   DOMAIN       28     83        PRO-RICH.
FT   DOMAIN      323    438        ZINC FINGERS.
FT   ZN_FING     323    347        C2H2-TYPE.
FT   ZN_FING     353    377        C2H2-TYPE.
FT   ZN_FING     383    405        C2H2-TYPE.
FT   ZN_FING     414    438        C2H2-TYPE.
FT   VARSPLIC    249    265        MISSING (IN ISOFORM 2 AND ISOFORM 3) (BY
FT                                 SIMILARITY).
FT   VARSPLIC    407    409        MISSING (IN ISOFORM 2 AND ISOFORM 4) (BY
FT                                 SIMILARITY).
SQ   SEQUENCE   449 AA;  49166 MW;  9C3E557B96F5A7B3 CRC64;


  Query Match             100.0%;  Score 56;  DB 1;  Length 449;
  Best Local Similarity   100.0%;  Pred. No. 0.012;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 CMTWNQMNL 9
          |||||||||
Db    235 CMTWNQMNL 243


RESULT   7
HIS5_LACLA
ID   HIS5_LACLA     STANDARD;      PRT;   202 AA.
AC   Q02132;
DT   01-JUL-1993 (Rel. 26, Created)
DT   01-JUL-1993 (Rel. 26, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   AMIDOTRANSFERASE HISH (EC 2.4.2.-).
GN   HISH.
OS   Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC   Lactococcus.
OX   NCBI_TaxID=1360;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NCDO 2118;
RX   MEDLINE=93015709; PubMed=1400209;
RA   Delorme C., Ehrlich S.D., Renault P.;
RT   "Histidine biosynthesis genes in Lactococcus lactis subsp. lactis.";
RL   J. Bacteriol. 174:6571-6579(1992).
```

```
CC   -!- FUNCTION: CATALYZES AN AMIDOTRANSFERASE REACTION THAT GENERATES
CC       IMIDAZOLE-GLYCEROL PHOSPHATE AND 5-AMINOIMIDAZOL-4-CARBOXAMIDE
CC       RIBONUCLEOTIDE, WHICH IS USED FOR PURINE SYNTHESIS.
CC   -!- PATHWAY: FIFTH STEP IN HISTIDINE BIOSYNTHETIC PATHWAY.
CC   -!- SUBCELLULAR LOCATION: CYTOPLASMIC (PROBABLE).
CC   -!- SIMILARITY: BELONGS TO THE HISH FAMILY.
CC   -!- SIMILARITY: CONTAINS 1 TYPE-1 GLUTAMINE AMIDOTRANSFERASE DOMAIN.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; U92974; AAB81907.1; -.
DR   PIR; I45734; I45734.
DR   InterPro; IPR000991; -.
DR   Pfam; PF00117; GATase; 1.
DR   PROSITE; PS00442; GATASE_TYPE_I; 1.
KW   Histidine biosynthesis; Transferase; Glutamine amidotransferase.
FT   ACT_SITE     81     81        BY SIMILARITY.
FT   ACT_SITE    183    183        BY SIMILARITY.
FT   ACT_SITE    185    185        BY SIMILARITY.
SQ   SEQUENCE   202 AA;  22440 MW;  E82909B78DB862CB CRC64;


  Query Match             66.1%;  Score 37;  DB 1;  Length 202;
  Best Local Similarity   75.0%;  Pred. No. 9.5;
  Matches    6;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      2 MTWNQMNL 9
          | |||:||
Db    119 MGWNQLNL 126


RESULT   8
COX2_ALBTU
ID   COX2_ALBTU     STANDARD;      PRT;   224 AA.
AC   Q09334;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   01-OCT-1996 (Rel. 34, Last annotation update)
DE   CYTOCHROME C OXIDASE POLYPEPTIDE II (EC 1.9.3.1).
GN   COII.
OS   Albinaria turrita (Door snail).
OG   Mitochondrion.
OC   Eukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata; Stylommatophora;
OC   Clausiliidae; Alopiinae; Albinaria.
OX   NCBI_TaxID=27820;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94275881; PubMed=8007005;
RA   Lecanidou R., Douris V., Rodakis G.C.;
RT   "Novel features of metazoan mtDNA revealed from sequence analysis of
RT   three mitochondrial DNA segments of the land snail Albinaria turrita
RT   (Gastropoda: Clausiliidae).";
RL   J. Mol. Evol. 38:369-382(1994).
CC   -!- FUNCTION: CYTOCHROME C OXIDASE IS THE COMPONENT OF THE RESPIRATORY
CC       CHAIN THAT CATALYZES THE REDUCTION OF OXYGEN TO WATER. SUBUNITS 1-
CC       3 FORM THE FUNCTIONAL CORE OF THE ENZYME COMPLEX. SUBUNIT 2
CC       TRANSFERS THE ELECTRONS FROM CYTOCHROME C VIA ITS BINUCLEAR COPPER
CC       A CENTER TO THE BIMETALLIC CENTER OF THE CATALYTIC SUBUNIT 1.
CC   -!- CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C + O(2) = 2 H(2)O +
CC       4 FERRICYTOCHROME C.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. MITOCHONDRIAL
CC       INNER MEMBRANE.
CC   -!- SIMILARITY: BELONGS TO THE CYTOCHROME C OXIDASE SUBUNIT 2 FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
```

```
DR   EMBL; M80218; AAA61299.1; JOINED.
DR   EMBL; M80219; AAA61299.1; JOINED.
DR   EMBL; M80220; AAA61299.1; JOINED.
DR   EMBL; M80221; AAA61299.1; JOINED.
DR   EMBL; M80228; AAA61299.1; JOINED.
DR   EMBL; M80229; AAA61299.1; JOINED.
DR   EMBL; M80231; AAA61299.1; JOINED.
DR   EMBL; M30393; AAA36810.1; -.
DR   PIR; A34673; A34673.
DR   PIR; S08273; S08273.
DR   TRANSFAC; T00899; -.
DR   MIM; 194070; -.
DR   MIM; 194080; -.
DR   MIM; 256370; -.
DR   InterPro; IPR000822; -.
DR   InterPro; IPR000976; -.
DR   Pfam; PF00096; zf-C2H2; 4.
DR   PRINTS; PR00048; ZINCFINGER.
DR   PRINTS; PR00049; WILMSTUMOUR.
DR   PROSITE; PS00028; ZINC_FINGER_C2H2_1; 4.
DR   PROSITE; PS50157; ZINC_FINGER_C2H2_2; 4.
KW   Zinc-finger; Metal-binding; DNA-binding; Repeat; Nuclear protein;
KW   Transcription regulation; Alternative splicing; Anti-oncogene;
KW   Disease mutation.
FT   DOMAIN       27     83       PRO-RICH.
FT   DOMAIN      323    438       ZINC FINGERS.
FT   ZN_FING     323    347       C2H2-TYPE.
FT   ZN_FING     353    377       C2H2-TYPE.
FT   ZN_FING     383    405       C2H2-TYPE.
FT   ZN_FING     414    438       C2H2-TYPE.
FT   VARSPLIC    250    266       MISSING (IN ISOFORM 2 AND ISOFORM 3).
FT   VARSPLIC    408    410       MISSING (IN ISOFORM 2 AND ISOFORM 4).
FT   VARIANT     181    181       P -> S (IN WT).
FT                                /FTId=VAR_007739.
FT   VARIANT     223    223       S -> N (IN WT).
FT                                /FTId=VAR_007740.
FT   VARIANT     253    253       G -> A (IN WT).
FT                                /FTId=VAR_007741.
FT   VARIANT     273    273       S -> G (IN MESOTHELIOMA).
FT                                /FTId=VAR_007742.
FT   VARIANT     330    330       C -> Y (IN DDS).
FT                                /FTId=VAR_007743.
FT   VARIANT     360    360       C -> G (IN DDS).
FT                                /FTId=VAR_007744.
FT   VARIANT     366    366       R -> C (IN WT).
FT                                /FTId=VAR_007745.
FT   VARIANT     366    366       R -> H (IN DDS).
FT                                /FTId=VAR_007746.
FT   VARIANT     373    373       H -> Q (IN DDS).

  Query Match             100.0%;  Score 56;  DB 1;  Length 449;
  Best Local Similarity   100.0%;  Pred. No. 0.012;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CMTWNQMNL 9
           |||||||||
Db     235 CMTWNQMNL 243


RESULT   5
WT1_MOUSE
ID   WT1_MOUSE      STANDARD;      PRT;   449 AA.
AC   P22561;
DT   01-AUG-1991 (Rel. 19, Created)
DT   01-AUG-1991 (Rel. 19, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   WILMS' TUMOR PROTEIN HOMOLOG.
GN   WT1 OR WT-1.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91141522; PubMed=1671709;
RA   Buckler A.J., Pelletier J., Haber D.A., Glaser T., Housman D.E.;
RT   "Isolation, characterization, and expression of the murine Wilms'
RT   tumor gene (WT1) during kidney development.";
RL   Mol. Cell. Biol. 11:1707-1712(1991).
CC   -!- FUNCTION: POTENTIAL ROLE IN TRANSCRIPTIONAL REGULATION. RECOGNIZES
CC       AND BINDS TO THE DNA SEQUENCE 5'-CGCCCCCGC-3'.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR.
CC   -!- ALTERNATIVE PRODUCTS: TWO ALTERNATIVE SPLICE SITES EXISTS.
CC   -!- TISSUE SPECIFICITY: KIDNEY.
CC   -!- DEVELOPMENTAL STAGE: EXPRESSED DURING KIDNEY DEVELOPMENT.
CC   -!- SIMILARITY: BELONGS TO THE EGR FAMILY OF C2H2-TYPE ZINC-FINGER
CC       PROTEINS.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; M55512; AAA40573.1; -.
DR   PIR; A39692; A39692.
DR   TRANSFAC; T02351; -.
DR   MGD; MGI:98968; Wt1.
DR   InterPro; IPR000822; -.
DR   InterPro; IPR000976; -.
DR   Pfam; PF00096; zf-C2H2; 4.
DR   PRINTS; PR00048; ZINCFINGER.
DR   PRINTS; PR00049; WILMSTUMOUR.
DR   PROSITE; PS00028; ZINC_FINGER_C2H2_1; 4.
DR   PROSITE; PS50157; ZINC_FINGER_C2H2_2; 4.
KW   Zinc-finger; Metal-binding; DNA-binding; Repeat; Nuclear protein;
KW   Transcription regulation; Alternative splicing; Anti-oncogene.
FT   DOMAIN       28     83       PRO-RICH.
FT   DOMAIN      323    438       ZINC FINGERS.
FT   ZN_FING     323    347       C2H2-TYPE.
FT   ZN_FING     353    377       C2H2-TYPE.
FT   ZN_FING     383    405       C2H2-TYPE.
FT   ZN_FING     414    438       C2H2-TYPE.
FT   VARSPLIC    250    266       MISSING (IN ISOFORM 2 AND ISOFORM 3).
FT   VARSPLIC    408    410       MISSING (IN ISOFORM 2 AND ISOFORM 4).
SQ   SEQUENCE   449 AA;  49246 MW;  962381E9C8D7A380 CRC64;

  Query Match             100.0%;  Score 56;  DB 1;  Length 449;
  Best Local Similarity   100.0%;  Pred. No. 0.012;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CMTWNQMNL 9
           |||||||||
Db     235 CMTWNQMNL 243


RESULT   6
WT1_PIG
ID   WT1_PIG        STANDARD;      PRT;   449 AA.
AC   O62651;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   WILMS' TUMOR PROTEIN HOMOLOG.
GN   WT1.
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX   NCBI_TaxID=9823;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=LWD; TISSUE=Kidney;
RX   MEDLINE=98267201; PubMed=9602131;
```

```
DE   WILMS' TUMOR PROTEIN (WT33).
GN   WT1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Fetal kidney;
RX   MEDLINE=90158822; PubMed=2154702;
RA   Gessler M., Poustka A., Cavenee W., Neve R.L., Orkin S.H.,
RA   Bruns G.A.P.;
RT   "Homozygous deletion in Wilms tumours of a zinc-finger gene
RT   identified by chromosome jumping.";
RL   Nature 343:774-778(1990).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Placenta;
RX   MEDLINE=92052142; PubMed=1658787;
RA   Haber D.A., Sohn R.L., Buckler A.J., Pelletier J., Call K.M.,
RA   Housman D.E.;
RT   "Alternative splicing and genomic structure of the Wilms tumor gene
RT   WT1.";
RL   Proc. Natl. Acad. Sci. U.S.A. 88:9618-9622(1991).
RN   [3]
RP   SEQUENCE OF 85-449 FROM N.A.
RX   MEDLINE=90150277; PubMed=2154335;
RA   Call K.M., Glaser T., Ito C.Y., Buckler A.J., Pelletier J.,
RA   Haber D.A., Rose E.A., Kral A., Yeger H., Lewis W.H., Jones C.,
RA   Housman D.E.;
RT   "Isolation and characterization of a zinc finger polypeptide gene at
RT   the human chromosome 11 Wilms' tumor locus.";
RL   Cell 60:509-520(1990).
RN   [4]
RP   IDENTIFICATION OF START CODON AND ALTERNATIVE SPLICE SITES.
RX   MEDLINE=91141522; PubMed=1671709;
RA   Buckler A.J., Pelletier J., Haber D.A., Glaser T., Housman D.E.;
RT   "Isolation, characterization, and expression of the murine Wilms'
RT   tumor gene (WT1) during kidney development.";
RL   Mol. Cell. Biol. 11:1707-1712(1991).
RN   [5]
RP   VARIANT WT CYS-366.
RX   MEDLINE=92279213; PubMed=1317572;
RA   Little M.H., Prosser J., Condie A., Smith P.J., van Heyningen V.,
RA   Hastie N.D.;
RT   "Zinc finger point mutations within the WT1 gene in Wilms tumor
RT   patients.";
RL   Proc. Natl. Acad. Sci. U.S.A. 89:4791-4795(1992).
RN   [6]
RP   VARIANTS DDS.
RX   MEDLINE=92005721; PubMed=1655284;
RA   Pelletier J., Bruening W., Kashtan C.E., Mauer S.M., Manivel J.C.,
RA   Striegel J.E., Houghton D.C., Junien C., Habib R., Fouser L.,
RA   Fine R.N., Silverman B.L., Haber D.A., Housman D.;
RT   "Germline mutations in the Wilms' tumor suppressor gene are
RT   associated with abnormal urogenital development in Denys-Drash
RT   syndrome.";
RL   Cell 67:437-447(1991).
RN   [7]
RP   VARIANTS DDS.
RX   MEDLINE=93265053; PubMed=1338906;
RA   Baird P.N., Santos A., Groves N., Jadresic L., Cowell J.K.;
RT   "Constitutional mutations in the WT1 gene in patients with
RT   Denys-Drash syndrome.";
RL   Hum. Mol. Genet. 1:301-305(1992).
RN   [8]
RP   VARIANTS DDS.
RX   MEDLINE=93271983; PubMed=8388765;
RA   Little M.H., Williamson K.A., Mannens M., Kelsey A., Gosden C.,
RA   Hastie N., van Heyningen V.;
RT   "Evidence that WT1 mutations in Denys-Drash syndrome patients may act
RT   in a dominant-negative fashion.";
RL   Hum. Mol. Genet. 2:259-264(1993).
RN   [9]
RP   VARIANT MESOTHELIOMA GLY-273.
RX   MEDLINE=94004972; PubMed=8401592;
RA   Park S., Schalling M., Bernard A., Maheswaran S., Shipley G.C.,
RA   Roberts D., Fletcher J., Shipman R., Rheinwald J., Demetri G.,
RA   Griffin J., Minden M., Housman D.E., Haber D.A.;
RT   "The Wilms tumour gene WT1 is expressed in murine mesoderm-derived
RT   tissues and mutated in a human mesothelioma.";
RL   Nat. Genet. 4:415-420(1993).
RN   [10]
RP   VARIANTS WT SER-181 AND ALA-253.
RX   MEDLINE=97268681; PubMed=9108089;
RA   Schumacher V., Schneider S., Figge A., Wildhardt G., Harms D.,
RA   Schmidt D., Weirich A., Ludwig R., Royer-Pokora B.;
RT   "Correlation of germ-line mutations and two-hit inactivation of the
RT   WT1 gene with Wilms tumors of stromal-predominant histology.";
RL   Proc. Natl. Acad. Sci. U.S.A. 94:3972-3977(1997).
RN   [11]
RP   VARIANTS DMS TYR-377; LEU-383 AND ASN-396.
RX   MEDLINE=98198341; PubMed=9529364;
RA   Jeanpierre C., Denamur E., Henry I., Cabanis M.-O., Luce S.,
RA   Cecille A., Elion J., Peuchmaur M., Loirat C., Niaudet P.,
RA   Gubler M.-C., Junien C.;
RT   "Identification of constitutional WT1 mutations, in patients with
RT   isolated diffuse mesangial sclerosis, and analysis of
RT   genotype/phenotype correlations by use of a computerized mutation
RT   database.";
RL   Am. J. Hum. Genet. 62:824-833(1998).
RN   [12]
RP   REVIEW.
RX   MEDLINE=92207913; PubMed=1313285;
RA   Haber D.A., Buckler A.J.;
RT   "WT1: a novel tumor suppressor gene inactivated in Wilms' tumor.";
RL   New Biol. 4:97-106(1992).
RN   [13]
RP   REVIEW.
RX   MEDLINE=93345769; PubMed=8393820;
RA   Rauscher F.J. III;
RT   "The WT1 Wilms tumor gene product: a developmentally regulated
RT   transcription factor in the kidney that functions as a tumor
RT   suppressor.";
RL   FASEB J. 7:896-903(1993).
CC   -!- FUNCTION: POTENTIAL ROLE IN TRANSCRIPTIONAL REGULATION. RECOGNIZES
CC       AND BINDS TO THE DNA SEQUENCE 5'-CGCCCCCGC-3'.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR.
CC   -!- ALTERNATIVE PRODUCTS: TWO ALTERNATIVE SPLICE SITES EXISTS.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN THE KIDNEY AND A SUBSET OF
CC       HEMATOPOIETIC CELLS.
CC   -!- DISEASE: WILM'S TUMOR (WT) IS AN EMBRYONAL MALIGNANCY OF THE
CC       KIDNEY THAT AFFECTS APPROXIMATELY 1 IN 10,000 INFANTS AND YOUNG
CC       CHILDREN. IT OCCURS BOTH IN SPORADIC AND HEREDITARY FORMS.
CC       INACTIVATION OF WT1 IS ONE OF THE CAUSES OF WILM'S TUMOR.
CC   -!- DISEASE: PATIENTS WITH DENYS-DRASH SYNDROME (DDS), WHICH IS
CC       CHARACTERIZED BY A TYPICAL NEPHROPATHY AND GENITAL ABNORMALITIES,
CC       HAVE DEFECTS IN THE WT1 GENE.
CC   -!- DISEASE: DEFECTS IN WT1 ARE ALSO A CAUSE OF DIFFUSE MESANGIAL
CC       SCLEROSIS (DMS), A FORM A DDS.
CC   -!- SIMILARITY: BELONGS TO THE EGR FAMILY OF C2H2-TYPE ZINC-FINGER
CC       PROTEINS.
CC   -!- DATABASE: NAME=WT1; NOTE=WT1 mutation database;
CC       WWW="http://www.umd.necker.fr:2003/".
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; X51630; CAA35956.1; ALT_INIT.
DR   EMBL; M80232; AAA61299.1; -.
DR   EMBL; M80217; AAA61299.1; JOINED.
```

```
  Best Local Similarity   100.0%;  Pred. No. 0.0064;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 CMTWNQMNL 9
          |||||||||
Db     42 CMTWNQMNL 50


RESULT   2
WT1_ALLMI
ID   WT1_ALLMI      STANDARD;      PRT;   288 AA.
AC   P50902;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   WILMS' TUMOR PROTEIN (FRAGMENT).
GN   WT1.
OS   Alligator mississippiensis (American alligator).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Crocodylidae; Alligatorinae; Alligator.
OX   NCBI_TaxID=8496;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96068905; PubMed=7478606;
RA   Kent J., Coriat A.M., Sharpe P.T., Hastie N., van Heyningen V.;
RT   "The evolution of WT1 sequence and expression pattern in the
RT   vertebrates.";
RL   Oncogene 11:1781-1792(1995).
CC   -!- FUNCTION: POTENTIAL ROLE IN TRANSCRIPTIONAL REGULATION. RECOGNIZES
CC       AND BINDS TO THE DNA SEQUENCE 5'-CGCCCCCGC-3'.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR.
CC   -!- ALTERNATIVE PRODUCTS: TWO ALTERNATIVE SPLICE SITES EXISTS (BY
CC       SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE EGR FAMILY OF C2H2-TYPE ZINC-FINGER
CC       PROTEINS.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; X85730; CAA59735.1; -.
DR   HSSP; P08046; 1A1L.
DR   InterPro; IPR000822; -.
DR   Pfam; PF00096; zf-C2H2; 4.
DR   PROSITE; PS00028; ZINC_FINGER_C2H2_1; 4.
DR   PROSITE; PS50157; ZINC_FINGER_C2H2_2; 4.
KW   Zinc-finger; Metal-binding; DNA-binding; Repeat; Nuclear protein;
KW   Transcription regulation; Alternative splicing.
FT   NON_TER       1      1
FT   ZN_FING     162    186       C2H2-TYPE.
FT   ZN_FING     192    216       C2H2-TYPE.
FT   ZN_FING     222    244       C2H2-TYPE.
FT   ZN_FING     253    277       C2H2-TYPE.
FT   VARSPLIC    245    247       MISSING (IN ISOFORM 2).
SQ   SEQUENCE   288 AA;  33111 MW;  33E26F7DBE7BFDDE CRC64;

  Query Match             100.0%;  Score 56;  DB 1;  Length 288;
  Best Local Similarity   100.0%;  Pred. No. 0.0076;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 CMTWNQMNL 9
          |||||||||
Db     89 CMTWNQMNL 97


RESULT   3
WT1_RAT
ID   WT1_RAT        STANDARD;      PRT;   448 AA.
AC   P49952;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   WILMS' TUMOR PROTEIN HOMOLOG.
GN   WT1 OR WT-1.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SPRAGUE-DAWLEY; TISSUE=Kidney;
RX   MEDLINE=93046155; PubMed=1330293;
RA   Sharma P.M., Yang X., Bowman M., Roberts V., Sukumar S.;
RT   "Molecular cloning of rat Wilms' tumor complementary DNA and a study
RT   of messenger RNA expression in the urogenital system and the brain.";
RL   Cancer Res. 52:6407-6412(1992).
CC   -!- FUNCTION: POTENTIAL ROLE IN TRANSCRIPTIONAL REGULATION. RECOGNIZES
CC       AND BINDS TO THE DNA SEQUENCE 5'-CGCCCCCGC-3'.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR.
CC   -!- ALTERNATIVE PRODUCTS: TWO ALTERNATIVE SPLICE SITES EXISTS.
CC   -!- TISSUE SPECIFICITY: KIDNEY (BY SIMILARITY).
CC   -!- DEVELOPMENTAL STAGE: EXPRESSED DURING KIDNEY DEVELOPMENT.
CC   -!- SIMILARITY: BELONGS TO THE EGR FAMILY OF C2H2-TYPE ZINC-FINGER
CC       PROTEINS.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; X69716; CAA49373.1; -.
DR   TRANSFAC; T02352; -.
DR   InterPro; IPR000822; -.
DR   InterPro; IPR000976; -.
DR   Pfam; PF00096; zf-C2H2; 4.
DR   PRINTS; PR00048; ZINCFINGER.
DR   PRINTS; PR00049; WILMSTUMOUR.
DR   PROSITE; PS00028; ZINC_FINGER_C2H2_1; 4.
DR   PROSITE; PS50157; ZINC_FINGER_C2H2_2; 4.
KW   Zinc-finger; Metal-binding; DNA-binding; Repeat; Nuclear protein;
KW   Transcription regulation; Alternative splicing; Anti-oncogene.
FT   DOMAIN       27     82       PRO-RICH.
FT   DOMAIN      322    437       ZINC FINGERS.
FT   ZN_FING     322    346       C2H2-TYPE.
FT   ZN_FING     352    376       C2H2-TYPE.
FT   ZN_FING     382    404       C2H2-TYPE.
FT   ZN_FING     413    437       C2H2-TYPE.
FT   VARSPLIC    249    265       MISSING (IN ISOFORM 2 AND ISOFORM 3).
FT   VARSPLIC    407    409       MISSING (IN ISOFORM 2 AND ISOFORM 4).
SQ   SEQUENCE   448 AA;  49193 MW;  329AC9AC1FF73F76 CRC64;

  Query Match             100.0%;  Score 56;  DB 1;  Length 448;
  Best Local Similarity   100.0%;  Pred. No. 0.011;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 CMTWNQMNL 9
          |||||||||
Db    234 CMTWNQMNL 242


RESULT   4
WT1_HUMAN
ID   WT1_HUMAN      STANDARD;      PRT;   449 AA.
AC   P19544;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-AUG-1991 (Rel. 19, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
```

```
                   GenCore version 4.5
            Copyright (c) 1993 - 2000  Compugen Ltd.


OM protein - protein search, using sw model

Run on:          October 23, 2001, 13:28:18 ; Search time 17.76 Seconds
                                              (without alignments)
                                              17.359 Million cell updates/sec

Title:           US-09-625-963-2
Perfect score:   56
Sequence:        1 CMTWNQMNL 9

Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched:        93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters:      93435

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       SwissProt_39:*

     Pred. No. is the number of results predicted by chance to have a
     score greater than or equal to the score of the result being printed,
     and is derived by analysis of the total score distribution.

                                 SUMMARIES
                 %
Result          Query
  No.   Score   Match Length DB  ID                    Description
-----------------------------------------------------------------------
    1      56   100.0    239  1  WT1_SMIMA             P49953 sminthopsis
    2      56   100.0    288  1  WT1_ALLMI             P50902 alligator m
    3      56   100.0    448  1  WT1_RAT               P49952 rattus norv
    4      56   100.0    449  1  WT1_HUMAN             P19544 homo sapien
    5      56   100.0    449  1  WT1_MOUSE             P22561 mus musculu
    6      56   100.0    449  1  WT1_PIG               O62651 sus scrofa
    7      37    66.1    202  1  HIS5_LACLA            Q02132 lactococcus
    8      36    64.3    224  1  COX2_ALBTU            Q09334 albinaria t
    9      36    64.3    245  1  YJF5_YEAST            P47044 saccharomyc
   10      36    64.3   1579  1  SSK2_YEAST            P53599 saccharomyc
   11      35    62.5    544  1  ESTA_DROPS            P25727 drosophila
   12      35    62.5    544  1  ESTP_DROME            P18167 drosophila
   13      35    62.5    545  1  ESTC_DROPS            P25725 drosophila
   14      35    62.5    712  1  PRTD_YEAST            P25375 saccharomyc
   15      35    62.5   1501  1  CDR3_CANAL            O42690 candida alb
   16      34    60.7    228  1  COX2_YPOMA            P98048 yponomeuta
   17      34    60.7    262  1  ROB2_AGRRH            P49409 agrobacteri
   18      34    60.7    385  1  YG1W_YEAST            P53230 saccharomyc
   19      34    60.7    445  1  DHE2_PORGI            Q03578 porphyromon
   20      34    60.7    616  1  SRE2_CAEEL            Q09273 caenorhabdi
   21      34    60.7    675  1  BROW_DROME            P12428 drosophila
   22      34    60.7    837  1  STA6_MOUSE            P52633 mus musculu
   23      34    60.7   1034  1  BGAL_BACME            O52847 bacillus me
   24      33    58.9    225  1  ATP6_LOCMI            P14569 locusta mig
   25      33    58.9    225  1  COX2_RHISA            O99819 rhipicephal
   26      33    58.9    287  1  FRHB_METJA            Q60341 methanococc
   27      33    58.9    304  1  VE4_HPV47             P22421 human papil
   28      33    58.9    310  1  YDHB_ECOLI            P37598 escherichia
   29      33    58.9    313  1  YQY2_CAEEL            Q09318 caenorhabdi
   30      33    58.9    452  1  RADA_LISMO            Q48761 listeria mo
   31      33    58.9    559  1  TPA_RAT               P19637 rattus norv
   32      33    58.9    562  1  TPA_HUMAN             P00750 homo sapien
   33      33    58.9    631  1  NTP1_VACCA            O57214 vaccinia vi
   34      33    58.9    631  1  NTP1_VACCC            P20637 vaccinia vi
   35      33    58.9    631  1  NTP1_VACCV            P05807 vaccinia vi
   36      33    58.9    631  1  NTP1_VARV             P33066 variola vir
   37      33    58.9    631  1  NTP1_YABAM            Q9qb93 yaba monkey
   38      33    58.9    632  1  NTP1_MYXVL            Q9q814 myxoma viru
   39      33    58.9    632  1  NTP1_SFVKA            Q9q8z2 shope fibro
   40      33    58.9    634  1  NTP1_MCV1             Q98267 molluscum c
   41      33    58.9    637  1  NTP1_FOWPV            O72907 fowlpox vir
   42      32    57.1    121  1  CRNF_LYMST            P91758 lymnaea sta
   43      32    57.1    202  1  LIF_RAT               P17777 rattus norv
   44      32    57.1    224  1  COX2_ALBCO            P48889 albinaria c
   45      32    57.1    224  1  COX2_EXERO            P29873 exeristes r

                                 ALIGNMENTS


RESULT   1
WT1_SMIMA
ID   WT1_SMIMA      STANDARD;      PRT;   239 AA.
AC   P49953;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   WILMS' TUMOR PROTEIN (FRAGMENT).
GN   WT1.
OS   Sminthopsis macroura (Stripe-faced dunnart).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Metatheria; Dasyuromorphia; Dasyuridae; Sminthopsis.
OX   NCBI_TaxID=9302;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Testis;
RX   MEDLINE=96068905; PubMed=7478606;
RA   Kent J., Coriat A.M., Sharpe P.T., Hastie N., van Heyningen V.;
RT   "The evolution of WT1 sequence and expression pattern in the
RT   vertebrates.";
RL   Oncogene 11:1781-1792(1995).
CC   -!- FUNCTION: POTENTIAL ROLE IN TRANSCRIPTIONAL REGULATION. RECOGNIZES
CC       AND BINDS TO THE DNA SEQUENCE 5'-CGCCCCCGC-3' (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: NUCLEAR.
CC   -!- ALTERNATIVE PRODUCTS: TWO ALTERNATIVE SPLICE SITES EXISTS (BY
CC       SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE EGR FAMILY OF C2H2-TYPE ZINC-FINGER
CC       PROTEINS.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; X85732; CAA59737.1; -.
DR   HSSP; P08046; 1A1L.
DR   InterPro; IPR000822; -.
DR   Pfam; PF00096; zf-C2H2; 4.
DR   PROSITE; PS00028; ZINC_FINGER_C2H2_1; 4.
DR   PROSITE; PS50157; ZINC_FINGER_C2H2_2; 4.
KW   Zinc-finger; Metal-binding; DNA-binding; Repeat; Nuclear protein;
KW   Transcription regulation; Alternative splicing; Anti-oncogene.
FT   NON_TER       1      1
FT   DOMAIN      113    228       ZINC FINGERS.
FT   ZN_FING     113    137       C2H2-TYPE.
FT   ZN_FING     143    167       C2H2-TYPE.
FT   ZN_FING     173    195       C2H2-TYPE.
FT   ZN_FING     204    228       C2H2-TYPE.
FT   VARSPLIC    198    200       MISSING (IN ISOFORM 2).
SQ   SEQUENCE   239 AA;  27793 MW;  6707678A7259A624 CRC64;


  Query Match             100.0%;  Score 56;  DB 1;  Length 239;
```

```
C;Accession: B34089
R;Collet, C.; Nielsen, K.M.; Russell, R.J.; Karl, M.; Oakeshott, J.G.; Richmond, R.C.
Mol. Biol. Evol. 7, 9-28, 1990
A;Title: Molecular analysis of duplicated esterase genes in Drosophila melanogaster.
A;Reference number: A34089; MUID:90136038
A;Accession: B34089
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-544 <COL>
A;Cross-references: GB:M33701; GB:M33780; GB:M33781; NID:g157377; PID:g157379
A;Note: the authors translated the codon CTG for residue 528 as Ala
C;Genetics:
A;Gene: FlyBase:Est-P
A;Cross-references: FlyBase:FBgn0000594
C;Superfamily: cholinesterase; cholinesterase homology
C;Keywords: carboxylic ester hydrolase


  Query Match             62.5%;  Score 35;  DB 2;  Length 544;
  Best Local Similarity   83.3%;  Pred. No. 1.4e+02;
  Matches    5;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 CMTWNQ 6
             || |||
Db      83 CMQWNQ 88


Search completed: October 23, 2001, 13:22:06
Job time: 119 sec
```

```
Qy          1 CMTWNQM 7
              || ||:|
Db         74 CMRWNRM 80


RESULT  11
S59801
protein kinase SSK2 (EC 2.7.1.-) - yeast (Saccharomyces cerevisiae)
N;Alternate names: protein N3276; protein YNR031c
C;Species: Saccharomyces cerevisiae
C;Date: 13-Jan-1996 #sequence_revision 23-Feb-1996 #text_change 24-Sep-1999
C;Accession: S59801; S63362
R;Maeda, T.; Takekawa, M.; Saito, H.
Science 269, 554-558, 1995
A;Title: Activation of yeast PBS2 MAPKK by MAPKKKs or by binding of an SH3-containing os
A;Reference number: S59801; MUID:95350642
A;Accession: S59801
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1-1579 <MAE>
A;Cross-references: EMBL:L41927; NID:g940329; PIDN:AAC41665.1; PID:g940330
R;Pohl, T.M.
submitted to the Protein Sequence Database, April 1996
A;Reference number: S63346
A;Accession: S63362
A;Molecule type: DNA
A;Residues: 1-1579 <POH>
A;Cross-references: EMBL:Z71646; NID:g1302526; PIDN:CAA96311.1; PID:e239826; PID:g130252
A;Experimental source: strain S288C
C;Genetics:
A;Gene: SGD:SSK2
A;Cross-references: SGD:S0005314; MIPS:YNR031c
A;Map position: 14R
C;Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolc
C;Keywords: ATP; phosphotransferase; protein kinase
F;1264-1558/Domain: protein kinase homology <KIN>
F;1272-1280/Region: protein kinase ATP-binding motif


  Query Match             64.3%;  Score 36;  DB 2;  Length 1579;
  Best Local Similarity   71.4%;  Pred. No. 2.6e+02;
  Matches    5;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 TWNQMNL 9
              |||::||
Db        609 TWNELNL 615


RESULT  12
T10542
hypothetical protein F3I3.40 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 02-Sep-2000
C;Accession: T10542
R;Bevan, M.; Van Der Schueren, J.; Chuang, Y.J.; Voet, M.; Robben, J.; Volckaert, G.; Ba
submitted to the Protein Sequence Database, June 1999
A;Reference number: Z17080
A;Accession: T10542
A;Molecule type: DNA
A;Residues: 1-2322 <BEV>
A;Cross-references: EMBL:AL080237; GSPDB:GN00062; ATSP:F3I3.40
A;Experimental source: cultivar Columbia; BAC clone F3I3
C;Genetics:
A;Gene: ATSP:F3I3.40
A;Map position: 4
A;Introns: 320/3; 775/3; 1366/1; 1774/1; 1832/2; 1944/3; 1982/1; 2059/3; 2141/1; 2233/3
C;Superfamily: RING finger homology
F;1914-1961/Domain: RING finger homology <RRN>

  Query Match             64.3%;  Score 36;  DB 2;  Length 2322;
  Best Local Similarity   62.5%;  Pred. No. 3.7e+02;
```

```
  Matches    5;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 CMTWNQMN 8
              | |||| :
Db         44 CTTWNQQH 51


RESULT  13
F84454
hypothetical protein At2g04100 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C;Accession: F84454
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon,
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487
A;Accession: F84454
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-480 <STO>
A;Cross-references: GB:AE002093; NID:g4734005; PIDN:AAD28682.1; GSPDB:GN00139
C;Genetics:
A;Gene: At2g04100
A;Map position: 2


  Query Match             63.4%;  Score 35.5;  DB 2;  Length 480;
  Best Local Similarity   60.0%;  Pred. No. 1e+02;
  Matches    6;  Conservative    3;  Mismatches    0;  Indels    1;  Gaps    1;

Qy          1 CMTWN-QMNL 9
              ||:|| |:|:
Db        466 CMSWNEQVNI 475


RESULT  14
T51405
ribosomal protein 3-like protein F14F8.140 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 18-Aug-2000 #sequence_revision 18-Aug-2000 #text_change 02-Sep-2000
C;Accession: T51405
R;Sato, S.; Nakamura, Y.; Kaneko, T.; Kato, T.; Asamizu, E.; Kotani, H.; Tabata, S.;
submitted to the Protein Sequence Database, August 2000
A;Reference number: Z25394
A;Accession: T51405
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-183 <SAT>
A;Cross-references: EMBL:AL391144
A;Experimental source: cultivar Columbia; BAC clone F14F8
C;Genetics:
A;Map position: 5
A;Introns: 89/3
A;Note: F14F8_140


  Query Match             62.5%;  Score 35;  DB 2;  Length 183;
  Best Local Similarity   83.3%;  Pred. No. 50;
  Matches    5;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 CMTWNQ 6
              | ||||
Db         46 CSTWNQ 51


RESULT  15
B34089
carboxylesterase (EC 3.1.1.1) P - fruit fly (Drosophila melanogaster)
C;Species: Drosophila melanogaster
C;Date: 01-Jun-1990 #sequence_revision 25-Sep-1992 #text_change 24-Nov-1999
```

```
  Matches    6;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 MTWNQMNL 9
           :|||| ||
Db     129 LTWNQANL 136


RESULT   6
I45734
imidazoleglycerol-phosphate synthase (EC 2.4.2.-) hisH - Lactococcus lactis subsp. lacti
C;Species: Lactococcus lactis subsp. lactis
C;Date: 24-Feb-1994 #sequence_revision 18-Nov-1994 #text_change 05-May-2000
C;Accession: I45734
R;Delorme, C.; Ehrlich, S.D.; Renault, P.
J. Bacteriol. 174, 6571-6579, 1992
A;Title: Histidine biosynthesis genes in Lactococcus lactis subsp. lactis.
A;Reference number: A45734; MUID:93015709
A;Contents: NCDO2118
A;Accession: I45734
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-202 <DEL>
A;Cross-references: GB:U92974; GB:M90760; GB:M90761; NID:g2565137; PIDN:AAB81907.1; PID:
A;Note: sequence extracted from NCBI backbone (NCBIN:115814, NCBIP:115831)
C;Superfamily: amidotransferase hisH; trpG homology
C;Keywords: glycosyltransferase; pentosyltransferase


  Query Match             66.1%;  Score 37;  DB 2;  Length 202;
  Best Local Similarity   75.0%;  Pred. No. 25;
  Matches    6;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 MTWNQMNL 9
           | |||:||
Db     119 MGWNQLNL 126


RESULT   7
D86776
amidotransferase (EC 2.4.2.-) [imported] - Lactococcus lactis subsp. lactis (strain IL14
C;Species: Lactococcus lactis subsp. lactis
C;Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 31-Mar-2001
C;Accession: D86776
R;Bolotin, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarme, K.; Weissenbach, J.; Ehrli
Genome Res. in press, 2001
A;Title: The complete genome sequence of the lactic acid bacterium.
A;Reference number: A86625
A;Accession: D86776
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-202 <STO>
A;Cross-references: GB:AE005176; NID:g12724181; PIDN:AAK05310.1; GSPDB:GN00146
A;Experimental source: strain IL1403
C;Genetics:
A;Gene: hisH
C;Superfamily: amidotransferase hisH; trpG homology
C;Keywords: glycosyltransferase; pentosyltransferase


  Query Match             66.1%;  Score 37;  DB 2;  Length 202;
  Best Local Similarity   75.0%;  Pred. No. 25;
  Matches    6;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 MTWNQMNL 9
           | |||:||
Db     119 MGWNQLNL 126


RESULT   8
B60407
monocyte adherence-induced protein 5 alpha - human
C;Species: Homo sapiens (man)
C;Date: 06-Nov-1992 #sequence_revision 06-Nov-1992 #text_change 03-May-1996
C;Accession: B60407
R;Sporn, S.A.; Eierman, D.F.; Johnson, C.E.; Morris, J.; Martin, G.; Ladner, M.; Hask
J. Immunol. 144, 4434-4441, 1990
A;Title: Monocyte adherence results in selective induction of novel genes sharing hom
A;Reference number: A60407; MUID:90257367
A;Accession: B60407
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-52 <SPO>
C;Superfamily: macrophage inflammatory protein


  Query Match             64.3%;  Score 36;  DB 2;  Length 52;
  Best Local Similarity   100.0%;  Pred. No. 10;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CMTWN 5
             |||||
Db      48 CMTWN 52


RESULT   9
S56827
conserved hypothetical protein YJL055w - yeast (Saccharomyces cerevisiae)
N;Alternate names: hypothetical protein J1148
C;Species: Saccharomyces cerevisiae
C;Date: 08-Jul-1995 #sequence_revision 08-Sep-1995 #text_change 08-Oct-1999
C;Accession: S56827
R;Pohl, T.M.; Aljinovic, G.
submitted to the Protein Sequence Database, September 1995
A;Reference number: S56793
A;Accession: S56827
A;Molecule type: DNA
A;Residues: 1-245 <TOV>
A;Cross-references: EMBL:Z49330; NID:g1008194; PIDN:CAA89346.1; PID:g1008195; MIPS:YJ
C;Genetics:
A;Map position: 10L
C;Superfamily: yeast conserved hypothetical protein YJL055w


  Query Match             64.3%;  Score 36;  DB 2;  Length 245;
  Best Local Similarity   44.4%;  Pred. No. 44;
  Matches    4;  Conservative    4;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 CMTWNQMNL 9
             |:||:|: :
Db     160 CITWSQLGI 168


RESULT  10
T20873
hypothetical protein F14B6.3 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T20873
R;White, S.
submitted to the EMBL Data Library, November 1996
A;Reference number: Z19338
A;Accession: T20873
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-366 <WIL>
A;Cross-references: EMBL:Z81502; PIDN:CAB04104.1; GSPDB:GN00019; CESP:F14B6.3
A;Experimental source: clone F14B6
C;Genetics:
A;Gene: CESP:F14B6.3
A;Map position: 1
A;Introns: 19/2; 84/3


  Query Match             64.3%;  Score 36;  DB 2;  Length 366;
  Best Local Similarity   71.4%;  Pred. No. 65;
  Matches    5;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;
```

```
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CMTWNQMNL 9
           |||||||||
Db     234 CMTWNQMNL 242


RESULT   3
A38080
Wilms tumor susceptibility protein WT1 - human
C;Species: Homo sapiens (man)
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 20-Jun-2000
C;Accession: A38080; S08273; A34673; I38504; I52811; I58315; A56411; S26286
R;Gessler, M.; Konig, A.; Bruns, G.A.
Genomics 12, 807-813, 1992
A;Title: The genomic organization and expression of the WT1 gene.
A;Reference number: A38080; MUID:92241883
A;Accession: A38080
A;Molecule type: DNA
A;Residues: 1-449 <GES1>
A;Cross-references: GB:X61631; GB:S99414; NID:g37981; PIDN:CAA43819.1; PID:g825731
A;Note: it is uncertain whether Met-1 is the initiator or whether translation is initiat
A;Note: sequence extracted from NCBI backbone (NCBIN:99414, NCBIN:99422, NCBIN:99479, NC
A;Note: the sequence in GenBank entry HSWTGEEX1, release 113.0, PIDN:CAA43819.1 differs
R;Gessler, M.; Poustka, A.; Cavenee, W.; Neve, R.L.; Orkin, S.H.; Bruns, G.A.P.
Nature 343, 774-778, 1990
A;Title: Homozygous deletion in Wilms tumours of a zinc-finger gene identified by chromo
A;Reference number: S08273; MUID:90158822
A;Accession: S08273
A;Molecule type: mRNA
A;Residues: 'SRQRPHPGALRNPTACPLPHFPPSLPPTHSPTHPPRAGTAAQAPGPRRLLAAILDFLLLQDPASTCVPEPASQHT
A;Cross-references: EMBL:X51630; NID:g37977; PIDN:CAA35956.1; PID:g37978
R;Call, K.M.; Glaser, T.; Ito, C.Y.; Buckler, A.J.; Pelletier, J.; Haber, D.A.; Rose, E.
Cell 60, 509-520, 1990
A;Title: Isolation and characterization of a zinc finger polypeptide gene at the human c
A;Reference number: A34673; MUID:90150277
A;Accession: A34673
A;Molecule type: mRNA
A;Residues: 85-249,267-364,'F',366-386,'T',388-407,411-449 <CAL>
A;Cross-references: GB:M30393; NID:g340381; PIDN:AAA36810.1; PID:g340382
R;Haber, D.A.; Sohn, R.L.; Buckler, A.J.; Pelletier, J.; Call, K.M.; Housman, D.E.
Proc. Natl. Acad. Sci. U.S.A. 88, 9618-9622, 1991
A;Title: Alternative splicing and genomic structure of the Wilms tumor gene WT1.
A;Reference number: A56411; MUID:92052142
A;Contents: annotation; alternative splicing
R;Phelan, S.A.; Lindberg, C.; Call, K.M.
Cell Growth Differ. 5, 677-686, 1994
A;Title: Wilms' tumor gene, WT1, mRNA is down-regulated during induction of erythroid an
A;Reference number: I38504; MUID:94368704
A;Accession: I38504
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-18 <PHE>
A;Cross-references: EMBL:U06486; NID:g473563; PIDN:AAA62865.1; PID:g458432
R;Pelletier, J.; Bruening, W.; Kashtan, C.E.; Mauer, S.M.; Manivel, J.C.; Striegel, J.E.
Cell 67, 437-447, 1991
A;Title: Germline mutations in the Wilms' tumor suppressor gene are associated with abno
A;Reference number: I52811; MUID:92005721
A;Accession: I52811
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 355-365,'H',367-377 <PEL>
A;Cross-references: GB:S61513; NID:g237599; PIDN:AAB20109.1; PID:g237600
A;Note: mutant form
R;Hamilton, T.B.; Barilla, K.C.; Romaniuk, P.J.
Nucleic Acids Res. 23, 277-284, 1995
A;Title: High affinity binding sites for the Wilms' tumour suppressor protein WT1.
A;Reference number: I58315; MUID:95166649
A;Accession: I58315
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 'MGHHHHHHHHHHSSGHIEGRHM',301-364,'F',366-386,'T',388-407,411-449 <HAM>
A;Cross-references: GB:S75264; NID:g896246; PIDN:AAB33443.1; PID:g896247
```

```
A;Note: this sequence is engineered
C;Genetics:
A;Gene: GDB:WT1
A;Cross-references: GDB:120496; OMIM:194070
A;Map position: 11p13-11p13
A;Introns: 148/1; 189/1; 223/2; 249/2; 266/2; 298/3; 349/1; 379/1; 410/1
A;Note: mRNA transcripts containing both alternatively spliced regions are the most a
C;Keywords: alternative splicing; DNA binding; kidney; tumor suppressor; zinc finger
F;1-449/Product: Wilms tumor susceptibility protein WT1, splice form 1 #status predic
F;1-407,411-449/Product: Wilms tumor susceptibility protein WT1, splice form 4 #statu
F;1-249,267-449/Product: Wilms tumor susceptibility protein WT1, splice form 3 #statu
F;1-249,267-407,411-449/Product: Wilms tumor susceptibility protein WT1, splice form


  Query Match             100.0%;  Score 56;  DB 2;  Length 449;
  Best Local Similarity   100.0%;  Pred. No. 0.032;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CMTWNQMNL 9
           |||||||||
Db     235 CMTWNQMNL 243


RESULT   4
A39692
Wilms' tumor protein analog, WT1 - mouse
C;Species: Mus musculus (house mouse)
C;Date: 30-Dec-1991 #sequence_revision 30-Dec-1991 #text_change 16-Feb-1997
C;Accession: A39692
R;Buckler, A.J.; Pelletier, J.; Haber, D.A.; Glaser, T.; Housman, D.E.
Mol. Cell. Biol. 11, 1707-1712, 1991
A;Title: Isolation, characterization, and expression of the murine Wilms' tumor gene
A;Reference number: A39692; MUID:91141522
A;Accession: A39692
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-449 <BUC>
A;Cross-references: GB:M55512
C;Keywords: alternative splicing; DNA binding; transcription regulation; tumor suppre


  Query Match             100.0%;  Score 56;  DB 2;  Length 449;
  Best Local Similarity   100.0%;  Pred. No. 0.032;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CMTWNQMNL 9
           |||||||||
Db     235 CMTWNQMNL 243


RESULT   5
G83933
dipeptidase BH2271 [imported] - Bacillus halodurans (strain C-125)
C;Species: Bacillus halodurans
C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 31-Dec-2000
C;Accession: G83933
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; H
Nucleic Acids Res. 28, 4317-4331, 2000
A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans a
A;Reference number: A83650; MUID:20263314
A;Accession: G83933
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-310 <STO>
A;Cross-references: GB:AP001515; GB:BA000004; NID:g10174886; PIDN:BAB05990.1; GSPDB:G
A;Experimental source: strain C-125
C;Genetics:
A;Gene: BH2271
C;Superfamily: membrane dipeptidase


  Query Match              67.9%;  Score 38;  DB 2;  Length 310;
  Best Local Similarity    75.0%;  Pred. No. 25;
```

```
                       GenCore version 4.5
              Copyright (c) 1993 - 2000  Compugen Ltd.

OM protein - protein search, using sw model

Run on:          October 23, 2001, 13:22:05 ; Search time 27.18 Seconds
                                              (without alignments)
                                              25.223 Million cell updates/sec

Title:           US-09-625-963-2
Perfect score:   56
Sequence:        1 CMTWNQMNL 9

Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched:        219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters:      219241

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       PIR_68:*
                1:  pir1:*
                2:  pir2:*
                3:  pir3:*
                4:  pir4:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                               SUMMARIES
                  %
Result           Query
  No.   Score   Match Length DB  ID                   Description
-------------------------------------------------------------------------------
    1      56   100.0    410  2  JC5046               Wilms' tumor suppr
    2      56   100.0    448  2  S33926               Wilms' tumor prote
    3      56   100.0    449  2  A38080               Wilms tumor suscep
    4      56   100.0    449  2  A39692               Wilms' tumor prote
    5      38    67.9    310  2  G83933               dipeptidase BH2271
    6      37    66.1    202  2  I45734               imidazoleglycerol-
    7      37    66.1    202  2  D86776               amidotransferase (
    8      36    64.3     52  2  B60407               monocyte adherence
    9      36    64.3    245  2  S56827               conserved hypothet
   10      36    64.3    366  2  T20873               hypothetical prote
   11      36    64.3   1579  2  S59801               protein kinase SSK
   12      36    64.3   2322  2  T10542               hypothetical prote
   13    35.5    63.4    480  2  F84454               hypothetical prote
   14      35    62.5    183  2  T51405               ribosomal protein
   15      35    62.5    544  2  B34089               carboxylesterase (
   16      35    62.5    712  2  S19387               saccharolysin (EC
   17      35    62.5    747  2  T33488               hypothetical prote
   18      35    62.5    849  1  T05181               S-receptor kinase
   19      35    62.5    850  2  T14450               serine/threonine k
   20      35    62.5   1124  2  T23061               hypothetical prote
   21      35    62.5   1369  2  S70713               protein-tyrosine k
   22      34    60.7     30  2  C49533               T-cell receptor al
   23      34    60.7    187  2  F72588               probable dTDP-4-de
   24      34    60.7    254  2  F82439               molybdopterin bios
   25      34    60.7    279  2  JC2106               rolB protein - Agr
   26      34    60.7    279  2  S52639               hypothetical prote
   27      34    60.7    299  2  T32982               hypothetical prote
   28      34    60.7    371  2  T00438               probable MYB famil
   29      34    60.7    371  2  T51636               myb-related transc
   30      34    60.7    385  2  S64340               hypothetical prote
   31      34    60.7    569  2  E85076               probable transposo
   32      34    60.7    616  2  T19873               hypothetical prote
   33      34    60.7    665  2  F84523               En/Spm-like transp
   34      34    60.7    675  1  FYFFB                brown protein - fr
   35      34    60.7    767  2  E85079               hypothetical prote
   36      34    60.7    837  2  I57557               DNA-Binding Protei
   37      34    60.7   1034  2  T30551               beta-galactosidase
   38      34    60.7   1034  2  T30574               beta-galactosidase
   39      33    58.9    217  2  T11382               cytochrome-c oxida
   40      33    58.9    225  2  T11155               cytochrome-c oxida
   41      33    58.9    225  2  T11471               H+-transporting AT
   42      33    58.9    231  2  B69756               ABC transporter (A
   43      33    58.9    237  2  I79268               traC protein - Esc
   44      33    58.9    254  2  I39721               ORF11 - Agrobacter
   45      33    58.9    277  2  B85354               hypothetical prote

                              ALIGNMENTS


RESULT   1
JC5046
Wilms' tumor suppressor protein - African clawed frog
N;Alternate names: WT1
C;Species: Xenopus laevis (African clawed frog)
C;Date: 31-Jan-1997 #sequence_revision 31-Jan-1997 #text_change 31-Jan-1997
C;Accession: JC5046
R;Semba, K.; Saito-Ueno, R.; Takayama, G.; Kondo, M.
Gene 175, 167-172, 1996
A;Title: cDNA cloning and its pronephros-specific expression of the Wilms' tumor supp
A;Reference number: JC5046; MUID:97074667
A;Contents: testis
A;Accession: JC5046
A;Molecule type: mRNA
A;Residues: 1-410 <SEM>
A;Cross-references: DDBJ:D82051
C;Comment: This protein is involved in kidney morphogenesis.
C;Genetics:
A;Gene: wt1


  Query Match             100.0%;  Score 56;  DB 2;  Length 410;
  Best Local Similarity   100.0%;  Pred. No. 0.029;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CMTWNQMNL 9
             |||||||||
Db     212 CMTWNQMNL 220


RESULT   2
S33926
Wilms' tumor protein WT1 - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 10-May-1996
C;Accession: S33926
R;Sharma, P.M.; Yang, X.; Bowman, M.; Roberts, V.; Sukumar, S.
Cancer Res. 52, 6407-6412, 1992
A;Title: Molecular cloning of rat Wilms' tumor complementary DNA and a study of messe
A;Reference number: S33926; MUID:93046155
A;Accession: S33926
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-448 <SHA>
A;Cross-references: EMBL:X69716
C;Genetics:
A;Gene: WT1
C;Keywords: tumor suppressor


  Query Match             100.0%;  Score 56;  DB 2;  Length 448;
  Best Local Similarity   100.0%;  Pred. No. 0.032;
```

```
;      NAME:  Parmelee, Steven W.
;      REGISTRATION NUMBER:  31,990
;      REFERENCE/DOCKET NUMBER:  14137-26-4PC
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 467-9600
;      TELEFAX:  (415) 543-5043
;  INFORMATION FOR SEQ ID NO:  52:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  10 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  unknown
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  peptide
PCT-US95-02121-52


  Query Match             43.1%;  Score 25;  DB 5;  Length 10;
  Best Local Similarity   50.0%;  Pred. No. 1.6e+02;
  Matches    5;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 HLMPFPGPLL 10
          :|:|  || |
Db      1 YLLPRRGPRL 10


Search completed: October 23, 2001, 13:35:07
Job time: 119 sec
```

```
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  No. 6140113e
US-09-063-276-14

  Query Match             46.6%;  Score 27;  DB 4;  Length 9;
  Best Local Similarity   66.7%;  Pred. No. 1.5e+05;
  Matches    6;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy         2 LMPFPGPLL 10
             | |||  ||
Db         1 LSPFPFDLL 9

RESULT  13
US-08-159-339A-813
; Sequence 813, Application US/08159339A
; Patent No. 6037135
;  GENERAL INFORMATION:
;    APPLICANT:  Kubo, Ralph T.
;    APPLICANT:  Grey, Howard M.
;    APPLICANT:  Sette, Alessandro
;    APPLICANT:  Celis, Esteban
;    TITLE OF INVENTION:  HLA Binding peptides and Their
;    TITLE OF INVENTION:  Uses
;    NUMBER OF SEQUENCES:  1254
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend and Crew LLP
;      STREET:  Two Embarcadero Center, Eighth Floor
;      CITY:  San Francisco
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94111-3834
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/159,339A
;      FILING DATE:  29-NOV-1993
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/926,666
;      FILING DATE:  07-AUG-1992
;      APPLICATION NUMBER:  US 08/027,746
;      FILING DATE:  05-MAR-1993
;      APPLICATION NUMBER:  US 08/103,396
;      FILING DATE:  06-AUG-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Weber, Ellen Lauver
;      REGISTRATION NUMBER:  32,762
;      REFERENCE/DOCKET NUMBER:  018623-005030US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 576-0200
;      TELEFAX:  (415) 576-0300
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  813:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-159-339A-813

  Query Match             44.8%;  Score 26;  DB 3;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 1.5e+05;
```

```
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         5 FPGP 8
             ||||
Db         5 FPGP 8

RESULT  14
US-08-916-935-10
; Sequence 10, Application US/08916935
; Patent No. 6193963
; GENERAL INFORMATION:
;  APPLICANT: Stern, Robert
;  APPLICANT: Frost, Gregory I.
;  APPLICANT: Csoka, Anthony
;  APPLICANT: Wong, Tim. M.
;  TITLE OF INVENTION: Human Plasma Hyaluronidase
;  FILE REFERENCE: 9076-088CIP
;  CURRENT APPLICATION NUMBER: US/08/916,935
;  CURRENT FILING DATE: 1997-08-21
;  EARLIER APPLICATION NUMBER: 08/733,360
;  EARLIER FILING DATE: 1996-10-17
;  NUMBER OF SEQ ID NOS: 11
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 10
;   LENGTH: 8
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-08-916-935-10

  Query Match             43.1%;  Score 25;  DB 4;  Length 8;
  Best Local Similarity   83.3%;  Pred. No. 1.5e+05;
  Matches    5;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         5 FPGPLL 10
             | ||||
Db         1 FRGPLL 6

RESULT  15
PCT-US95-02121-52
; Sequence 52, Application PC/TUS9502121
;  GENERAL INFORMATION:
;    APPLICANT:
;    TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR ELICITING
;    TITLE OF INVENTION:  CTL IMMUNITY
;    NUMBER OF SEQUENCES:  153
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/02121
;      FILING DATE:  16-FEB-1995
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/197,484
;      FILING DATE:  16-FEB-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/935,811
;      FILING DATE:  26-AUG-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/874,491
;      FILING DATE:  27-APR-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/827,682
;      FILING DATE:  29-JAN-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/749,568
;      FILING DATE:  26-AUG-1991
;    ATTORNEY/AGENT INFORMATION:
```

```
;    APPLICANT:  STEVEN A.
;    TITLE OF INVENTION:  MELANOMA ANTIGENS AND
;    TITLE OF INVENTION:  THEIR USE IN DIAGNOSTIC AND THERAPEUTIC
;    TITLE OF INVENTION:  METHODS
;    NUMBER OF SEQUENCES:  126
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MORGAN & FINNEGAN, L.L.P.
;      STREET:  345 PARK AVENUE
;      CITY:  NEW YORK
;      STATE:  NEW YORK
;      COUNTRY:  USA
;      ZIP:  10154
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  FLOPPY DISK
;      COMPUTER:  IBM PC COMPATIBLE
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  ASCII
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/417,174
;      FILING DATE:  05-APR-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/231,565
;      FILING DATE:  22-APR-1994
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  CAROL M. GRUPPI
;      REGISTRATION NUMBER:  37,341
;      REFERENCE/DOCKET NUMBER:  2026-4124US1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 758-4800
;      TELEFAX:  (212) 751-6849
;      TELEX:  421792
;  INFORMATION FOR SEQ ID NO:  119:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9
;      TYPE:  amino acid
;      STRANDEDNESS:  Unknown
;      TOPOLOGY:  Unknown
;    MOLECULE TYPE:  Peptide
US-08-417-174-119


  Query Match             46.6%;  Score 27;  DB 2;  Length 9;
  Best Local Similarity   80.0%;  Pred. No. 1.5e+05;
  Matches    4;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       5 FPGPL 9
           ||||:
Db       3 FPGPV 7


RESULT  11
US-08-828-712-14
; Sequence 14, Application US/08828712
; Patent No. 6015884
;  GENERAL INFORMATION:
;    APPLICANT:  Schneck, Jonathan P.
;    APPLICANT:  O'Herrin, Sean
;    TITLE OF INVENTION:  Soluble Divalent and Multivalent
;    TITLE OF INVENTION:  Heterodimeric Analogs of Proteins
;    NUMBER OF SEQUENCES:  20
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Banner & Witcoff, Ltd.
;      STREET:  1001 G Street, N.W.
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  U.S.A.
;      ZIP:  20001
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
```

```
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/828,712
;      FILING DATE:  28-MAR-1997
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Kagan, Sarah A.
;      REGISTRATION NUMBER:  32,141
;      REFERENCE/DOCKET NUMBER:  01107.73713
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)508-9100
;      TELEFAX:  (202)508-9299
;  INFORMATION FOR SEQ ID NO:  14:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    HYPOTHETICAL:  NO
;    ORIGINAL SOURCE:
;      INDIVIDUAL ISOLATE:  peptide SL9
US-08-828-712-14


  Query Match             46.6%;  Score 27;  DB 3;  Length 9;
  Best Local Similarity   66.7%;  Pred. No. 1.5e+05;
  Matches    6;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 LMPFPGPLL 10
           | |||  ||
Db       1 LSPFPFDLL 9


RESULT  12
US-09-063-276-14
; Sequence 14, Application US/09063276
; Patent No. 6140113
;  GENERAL INFORMATION:
;    APPLICANT:  Schneck, Jonathan
;    APPLICANT:  O'Herrin, Sean
;    TITLE OF INVENTION:  Molecular Complexes Which
;    TITLE OF INVENTION:  Modify Immune Responses
;    NUMBER OF SEQUENCES:  20
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Banner & Witcoff
;      STREET:  1001 G Street, NW
;      CITY:  Washington
;      STATE:  DC
;      COUNTRY:  USA
;      ZIP:  20001
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/063,276
;      FILING DATE:  21-APR-1998
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/828,712
;      FILING DATE:  28-MAR-1997
;      APPLICATION NUMBER:  60/014,367
;      FILING DATE:  28-MAR-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Kagan, Sarah A
;      REGISTRATION NUMBER:  32141
;      REFERENCE/DOCKET NUMBER:  01107.74154
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-508-9100
;      TELEFAX:  202-508-9299
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  14:
```

```
;    APPLICANT:  Hammond, David J.
;    TITLE OF INVENTION:  Peptides Which Bind to Prothrombin and
;    TITLE OF INVENTION:  Thrombin
;    NUMBER OF SEQUENCES:  21
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Bayer Corporation
;      STREET:  800 Dwight Way
;      STREET:  P. O. Box 1986
;      CITY:  Berkeley
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94701-1986
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.50 inch, 1.44Mb Storage
;      COMPUTER:  IBM
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  WordPerfect 6.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/672,805
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Giblin, James A.
;      REGISTRATION NUMBER:  25772
;      REFERENCE/DOCKET NUMBER:  MSB-7236
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (510)705-7910
;      TELEFAX:  (510)705-7904
;  INFORMATION FOR SEQ ID NO:  19:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  7
;      TYPE:  amino acid
;      STRANDEDNESS:  single strand
;      TOPOLOGY:  linear
;    MOLECULE TYPE:
;      DESCRIPTION:  peptide
US-08-672-805-19


  Query Match             50.0%;  Score 29;  DB 2;  Length 7;
  Best Local Similarity   83.3%;  Pred. No. 1.5e+05;
  Matches    5;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       5 FPGPLL 10
           |||| |
Db       2 FPGPYL 7


RESULT   8
US-09-001-984C-33
; Sequence 33, Application US/09001984C
; Patent No. 6245331
; GENERAL INFORMATION:
;  APPLICANT:  Laal, Suman
;  APPLICANT:  Zolla-Pazner, Susan
;  APPLICANT:  Belisle, John T
;  TITLE OF INVENTION:  EARLY DETECTION OF MYCOBACTERIAL DISEASE
;  FILE REFERENCE:  NYU-011
;  CURRENT APPLICATION NUMBER: US/09/001,984C
;  CURRENT FILING DATE:  1997-12-31
;  PRIOR APPLICATION NUMBER:  60/034,003
;  PRIOR FILING DATE:  1996-12-31
;  NUMBER OF SEQ ID NOS:  106
;  SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 33
;   LENGTH: 8
;   TYPE: PRT
;   ORGANISM: Mycobacterium tuberculosis strain H37Rv
US-09-001-984C-33
```

```
  Query Match             48.3%;  Score 28;  DB 4;  Length 8;
  Best Local Similarity   80.0%;  Pred. No. 1.5e+05;
  Matches    4;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3 MPFPG 7
           ||:||
Db       2 MPYPG 6


RESULT   9
US-08-417-174-111
; Sequence 111, Application US/08417174
; Patent No. 5844075
;  GENERAL INFORMATION:
;    APPLICANT:  KAWAKAMI, YUTAKA; ROSENBERG,
;    APPLICANT:  STEVEN A.
;    TITLE OF INVENTION:  MELANOMA ANTIGENS AND
;    TITLE OF INVENTION:  THEIR USE IN DIAGNOSTIC AND THERAPEUTIC
;    TITLE OF INVENTION:  METHODS
;    NUMBER OF SEQUENCES:  126
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MORGAN & FINNEGAN, L.L.P.
;      STREET:  345 PARK AVENUE
;      CITY:  NEW YORK
;      STATE:  NEW YORK
;      COUNTRY:  USA
;      ZIP:  10154
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  FLOPPY DISK
;      COMPUTER:  IBM PC COMPATIBLE
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  ASCII
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/417,174
;      FILING DATE:  05-APR-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/231,565
;      FILING DATE:  22-APR-1994
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  CAROL M. GRUPPI
;      REGISTRATION NUMBER:  37,341
;      REFERENCE/DOCKET NUMBER:  2026-4124US1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 758-4800
;      TELEFAX:  (212) 751-6849
;      TELEX:  421792
;  INFORMATION FOR SEQ ID NO:  111:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9
;      TYPE:  amino acid
;      STRANDEDNESS:  Unknown
;      TOPOLOGY:  Unknown
;    MOLECULE TYPE:  Peptide
US-08-417-174-111


  Query Match             46.6%;  Score 27;  DB 2;  Length 9;
  Best Local Similarity   80.0%;  Pred. No. 1.5e+05;
  Matches    4;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       5 FPGPL 9
           ||||:
Db       3 FPGPV 7


RESULT  10
US-08-417-174-119
; Sequence 119, Application US/08417174
; Patent No. 5844075
;  GENERAL INFORMATION:
;    APPLICANT:  KAWAKAMI, YUTAKA; ROSENBERG,
```

```
;    HYPOTHETICAL:  NO
;    FRAGMENT TYPE:  internal
PCT-US95-03934A-8

  Query Match             56.9%;  Score 33;  DB 5;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 1.5e+05;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      4 PFPGP 8
          |||||
Db      1 PFPGP 5


RESULT   5
US-08-159-339A-799
; Sequence 799, Application US/08159339A
; Patent No. 6037135
;  GENERAL INFORMATION:
;    APPLICANT:  Kubo, Ralph T.
;    APPLICANT:  Grey, Howard M.
;    APPLICANT:  Sette, Alessandro
;    APPLICANT:  Celis, Esteban
;    TITLE OF INVENTION:  HLA Binding peptides and Their
;    TITLE OF INVENTION:  Uses
;    NUMBER OF SEQUENCES:  1254
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend and Crew LLP
;      STREET:  Two Embarcadero Center, Eighth Floor
;      CITY:  San Francisco
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94111-3834
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/159,339A
;      FILING DATE:  29-NOV-1993
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/926,666
;      FILING DATE:  07-AUG-1992
;      APPLICATION NUMBER:  US 08/027,746
;      FILING DATE:  05-MAR-1993
;      APPLICATION NUMBER:  US 08/103,396
;      FILING DATE:  06-AUG-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Weber, Ellen Lauver
;      REGISTRATION NUMBER:  32,762
;      REFERENCE/DOCKET NUMBER:  018623-005030US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 576-0200
;      TELEFAX:  (415) 576-0300
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  799:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-159-339A-799


  Query Match             51.7%;  Score 30;  DB 3;  Length 9;
  Best Local Similarity   71.4%;  Pred. No. 1.5e+05;
  Matches    5;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      2 LMPFPGP 8
          :| ||||
Db      2 VMGFPGP 8


RESULT   6
US-08-159-339A-837
; Sequence 837, Application US/08159339A
; Patent No. 6037135
;  GENERAL INFORMATION:
;    APPLICANT:  Kubo, Ralph T.
;    APPLICANT:  Grey, Howard M.
;    APPLICANT:  Sette, Alessandro
;    APPLICANT:  Celis, Esteban
;    TITLE OF INVENTION:  HLA Binding peptides and Their
;    TITLE OF INVENTION:  Uses
;    NUMBER OF SEQUENCES:  1254
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend and Crew LLP
;      STREET:  Two Embarcadero Center, Eighth Floor
;      CITY:  San Francisco
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94111-3834
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/159,339A
;      FILING DATE:  29-NOV-1993
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/926,666
;      FILING DATE:  07-AUG-1992
;      APPLICATION NUMBER:  US 08/027,746
;      FILING DATE:  05-MAR-1993
;      APPLICATION NUMBER:  US 08/103,396
;      FILING DATE:  06-AUG-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Weber, Ellen Lauver
;      REGISTRATION NUMBER:  32,762
;      REFERENCE/DOCKET NUMBER:  018623-005030US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 576-0200
;      TELEFAX:  (415) 576-0300
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  837:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  10 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-159-339A-837


  Query Match             51.7%;  Score 30;  DB 3;  Length 10;
  Best Local Similarity   71.4%;  Pred. No. 25;
  Matches    5;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      2 LMPFPGP 8
          :| ||||
Db      3 VMGFPGP 9


RESULT   7
US-08-672-805-19
; Sequence 19, Application US/08672805
; Patent No. 5831003
;  GENERAL INFORMATION:
;    APPLICANT:  Baumbach, George A.,
;    APPLICANT:  Buettner, Joseph A.,
;    APPLICANT:  Dadd, Christopher A.,
```

```
  Query Match             56.9%;  Score 33;  DB 1;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 1.5e+05;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       4 PFPGP 8
           |||||
Db       1 PFPGP 5


RESULT   2
US-08-224-917-8
; Sequence 8, Application US/08224917
; Patent No. 5965350
;  GENERAL INFORMATION:
;    APPLICANT:  Lou, Lillian Lien-Li
;    APPLICANT:  Barnett, Jimmy Wayne
;    TITLE OF INVENTION:  Cloning and Expression of Human GMP
;    TITLE OF INVENTION:  Synthetase, its use in Screening for Inhibitors of Human
;    TITLE OF INVENTION:  GMP Synthetase and Inhibitors of Human GMP Synthetase
;    NUMBER OF SEQUENCES:  11
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Syntex (USA) Inc.
;      STREET:  3401 Hillview Avenue
;      CITY:  Palo Alto
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/224,917
;      FILING DATE:  08-APR-1994
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Peries, Rohan
;      REGISTRATION NUMBER:  35,752
;      REFERENCE/DOCKET NUMBER:  28060
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415)-852-1698
;      TELEFAX:  (415)-496-3529
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    HYPOTHETICAL:  NO
;    FRAGMENT TYPE:  internal
US-08-224-917-8


  Query Match             56.9%;  Score 33;  DB 2;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 1.5e+05;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       4 PFPGP 8
           |||||
Db       1 PFPGP 5


RESULT   3
US-08-914-853-8
; Sequence 8, Application US/08914853
; Patent No. 5998186
;  GENERAL INFORMATION:
;    APPLICANT:  Lou, Lillian Lien-Li
;    APPLICANT:  Barnett, Jimmy Wayne
;    TITLE OF INVENTION:  Cloning and Expression of Human GMP
;    TITLE OF INVENTION:  Synthetase, its use in Screening for Inhibitors of
;    TITLE OF INVENTION:  Human
;    TITLE OF INVENTION:  GMP Synthetase and Inhibitors of Human GMP Synthetase
;    NUMBER OF SEQUENCES:  11
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Syntex (USA) Inc.
;      STREET:  3401 Hillview Avenue
;      CITY:  Palo Alto
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/914,853
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/461,489
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Peries, Rohan
;      REGISTRATION NUMBER:  35,752
;      REFERENCE/DOCKET NUMBER:  28060
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415)-852-1698
;      TELEFAX:  (415)-496-3529
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    HYPOTHETICAL:  NO
;    FRAGMENT TYPE:  internal
US-08-914-853-8


  Query Match             56.9%;  Score 33;  DB 2;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 1.5e+05;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       4 PFPGP 8
           |||||
Db       1 PFPGP 5


RESULT   4
PCT-US95-03934A-8
; Sequence 8, Application PC/TUS9503934A
;  GENERAL INFORMATION:
;    APPLICANT:  Syntex (USA) Inc.
;    TITLE OF INVENTION:  Cloning and Expression of Human GMP
;    TITLE OF INVENTION:  Synthetase, its use in Screening for Inhibitors of Human
;    TITLE OF INVENTION:  GMP Synthetase and Inhibitors of Human GMP Synthetase
;    NUMBER OF SEQUENCES:  11
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/03934A
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
```

```
                    GenCore version 4.5
           Copyright (c) 1993 - 2000  Compugen Ltd.


OM protein - protein search, using sw model

Run on:        October 23, 2001, 13:33:08 ; Search time 12.16 Seconds
                                            (without alignments)
                                            16.933 Million cell updates/sec

Title:         US-09-625-963-3
Perfect score: 58
Sequence:      1 HLMPFPGPLL 10

Scoring table: BLOSUM62
               Gapop 10.0 , Gapext 0.5

Searched:      197339 seqs, 20590346 residues

Total number of hits satisfying chosen parameters:      59481

Minimum DB seq length: 0
Maximum DB seq length: 10

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       Issued_Patents_AA:*
               1:  /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
               2:  /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
               3:  /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
               4:  /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
               5:  /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
               6:  /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                                SUMMARIES
               %
Result       Query
  No.  Score Match Length DB ID                        Description
--------------------------------------------------------------------------
    1     33  56.9      9  1  US-08-463-620-8            Sequence 8, Appli
    2     33  56.9      9  2  US-08-224-917-8            Sequence 8, Appli
    3     33  56.9      9  2  US-08-914-853-8            Sequence 8, Appli
    4     33  56.9      9  5  PCT-US95-03934A-8          Sequence 8, Appli
    5     30  51.7      9  3  US-08-159-339A-799         Sequence 799, App
    6     30  51.7     10  3  US-08-159-339A-837         Sequence 837, App
    7     29  50.0      7  2  US-08-672-805-19           Sequence 19, Appl
    8     28  48.3      8  4  US-09-001-984C-33          Sequence 33, Appl
    9     27  46.6      9  2  US-08-417-174-111          Sequence 111, App
   10     27  46.6      9  2  US-08-417-174-119          Sequence 119, App
   11     27  46.6      9  3  US-08-828-712-14           Sequence 14, Appl
   12     27  46.6      9  4  US-09-063-276-14           Sequence 14, Appl
   13     26  44.8      9  3  US-08-159-339A-813         Sequence 813, App
   14     25  43.1      8  4  US-08-916-935-10           Sequence 10, Appl
   15     25  43.1     10  5  PCT-US95-02121-52          Sequence 52, Appl
   16     25  43.1     10  5  PCT-US95-02121-131         Sequence 131, App
   17     24  41.4      6  1  US-08-244-701B-39          Sequence 39, Appl
   18     24  41.4      8  5  PCT-US91-09152-7           Sequence 7, Appli
   19     24  41.4      9  2  US-08-417-174-109          Sequence 109, App
   20     24  41.4      9  2  US-08-417-174-118          Sequence 118, App
   21     24  41.4      9  3  US-08-993-165-29           Sequence 29, Appl
   22     23  39.7      7  2  US-08-500-635A-10          Sequence 10, Appl
   23     23  39.7      8  3  US-08-828-712-12           Sequence 12, Appl
   24     23  39.7      8  3  US-08-993-165-25           Sequence 25, Appl
   25     23  39.7      8  3  US-08-993-165-28           Sequence 28, Appl
   26     23  39.7      8  3  US-08-993-165-30           Sequence 30, Appl
   27     23  39.7      8  4  US-09-063-276-12           Sequence 12, Appl
   28     23  39.7      8  4  US-08-444-818-263          Sequence 263, App
   29     23  39.7      8  5  PCT-US93-11703-73          Sequence 73, Appl
   30     23  39.7      9  2  US-08-146-028-223          Sequence 223, App
   31     23  39.7      9  2  US-08-146-028-224          Sequence 224, App
   32     23  39.7      9  2  US-08-146-028-225          Sequence 225, App
   33     23  39.7      9  3  US-08-828-712-13           Sequence 13, Appl
   34     23  39.7      9  4  US-09-063-276-13           Sequence 13, Appl
   35     23  39.7      9  4  US-08-723-425A-223         Sequence 223, App
   36     23  39.7      9  4  US-08-723-425A-224         Sequence 224, App
   37     23  39.7      9  4  US-08-723-425A-225         Sequence 225, App
   38     23  39.7      9  4  US-09-112-206-223          Sequence 223, App
   39     23  39.7      9  4  US-09-112-206-224          Sequence 224, App
   40     23  39.7      9  4  US-09-112-206-225          Sequence 225, App
   41     22  37.9      8  3  US-08-159-339A-777         Sequence 777, App
   42     22  37.9      8  3  US-08-159-339A-1226        Sequence 1226, Ap
   43     22  37.9      8  5  PCT-US91-09152-9           Sequence 9, Appli
   44     22  37.9      9  1  US-08-269-441A-4           Sequence 4, Appli
   45     22  37.9      9  2  US-08-146-028-443          Sequence 443, App


                                ALIGNMENTS


RESULT   1
US-08-463-620-8
; Sequence 8, Application US/08463620
; Patent No. 5789216
;  GENERAL INFORMATION:
;    APPLICANT:  Lou, Lillian Lien-Li
;    APPLICANT:  Barnett, Jimmy Wayne
;    TITLE OF INVENTION:  Cloning and Expression of Human GMP
;    TITLE OF INVENTION:  Synthetase, its use in Screening for Inhibitors of
;    TITLE OF INVENTION:  Human
;    TITLE OF INVENTION:  GMP Synthetase and Inhibitors of Human GMP Synthetase
;    NUMBER OF SEQUENCES:  11
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Syntex (USA) Inc.
;      STREET:  3401 Hillview Avenue
;      CITY:  Palo Alto
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/463,620
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/224,917
;      FILING DATE:  08-APR-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Peries, Rohan
;      REGISTRATION NUMBER:  35,752
;      REFERENCE/DOCKET NUMBER:  28060
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415)-852-1698
;      TELEFAX:  (415)-496-3529
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    HYPOTHETICAL:  NO
;    FRAGMENT TYPE:  internal
US-08-463-620-8
```

```
RESULT  13
US-09-790-812A-211
; Sequence 211, Application US/09790812A
; GENERAL INFORMATION:
;  APPLICANT: De Leys, Robert
;  TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES CORRESPONDING
;  TITLE OF INVENTION:  TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR USE IN
;  TITLE OF INVENTION:  A PROCESS FOR DETERMINATION OF ANTIBODIES OF
;  TITLE OF INVENTION:  BIOTINYLATED PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT
;  TITLE OF INVENTION:  EPITOPES, A PROCESS FOR PREPARING THEM AND COMPOSITIONS
;  TITLE OF INVENTION:  CONTAINING THEM
;  FILE REFERENCE: 2752-17
;  CURRENT APPLICATION NUMBER: US/09/790,812A
;  CURRENT FILING DATE:  2001-02-23
;  PRIOR APPLICATION NUMBER: 09/576,824
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 08/723,425
;  PRIOR FILING DATE: 1996-09-30
;  PRIOR APPLICATION NUMBER: 09/146,028
;  PRIOR FILING DATE: 1993-11-22
;  PRIOR APPLICATION NUMBER: PCT/EP93/00517
;  PRIOR FILING DATE: 1993-03-08
;  PRIOR APPLICATION NUMBER: EP 92400598.6
;  PRIOR FILING DATE: 1992-03-06
;  NUMBER OF SEQ ID NOS: 600
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 211
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Hepatitis C virus
US-09-790-812A-211


  Query Match            39.7%;  Score 23;  DB 5;  Length 9;
  Best Local Similarity   50.0%;  Pred. No. 2.9e+05;
  Matches    4;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 HLMPFPGP 8
           :|:|  ||
Db       2 YLLPRRGP 9

RESULT  14
US-09-790-812A-212
; Sequence 212, Application US/09790812A
; GENERAL INFORMATION:
;  APPLICANT: De Leys, Robert
;  TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES CORRESPONDING
;  TITLE OF INVENTION:  TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR USE IN
;  TITLE OF INVENTION:  A PROCESS FOR DETERMINATION OF ANTIBODIES OF
;  TITLE OF INVENTION:  BIOTINYLATED PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT
;  TITLE OF INVENTION:  EPITOPES, A PROCESS FOR PREPARING THEM AND COMPOSITIONS
;  TITLE OF INVENTION:  CONTAINING THEM
;  FILE REFERENCE: 2752-17
;  CURRENT APPLICATION NUMBER: US/09/790,812A
;  CURRENT FILING DATE:  2001-02-23
;  PRIOR APPLICATION NUMBER: 09/576,824
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 08/723,425
;  PRIOR FILING DATE: 1996-09-30
;  PRIOR APPLICATION NUMBER: 09/146,028
;  PRIOR FILING DATE: 1993-11-22
;  PRIOR APPLICATION NUMBER: PCT/EP93/00517
;  PRIOR FILING DATE: 1993-03-08
;  PRIOR APPLICATION NUMBER: EP 92400598.6
;  PRIOR FILING DATE: 1992-03-06
;  NUMBER OF SEQ ID NOS: 600
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 212
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Hepatitis C virus
US-09-790-812A-212

  Query Match            39.7%;  Score 23;  DB 5;  Length 9;
  Best Local Similarity   50.0%;  Pred. No. 2.9e+05;
  Matches    4;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 HLMPFPGP 8
           :|:|  ||
Db       1 YLLPRRGP 8

RESULT  15
US-09-790-812A-213
; Sequence 213, Application US/09790812A
; GENERAL INFORMATION:
;  APPLICANT: De Leys, Robert
;  TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES CORRESPONDING
;  TITLE OF INVENTION:  TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR USE IN
;  TITLE OF INVENTION:  A PROCESS FOR DETERMINATION OF ANTIBODIES OF
;  TITLE OF INVENTION:  BIOTINYLATED PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORT
;  TITLE OF INVENTION:  EPITOPES, A PROCESS FOR PREPARING THEM AND COMPOSITIONS
;  TITLE OF INVENTION:  CONTAINING THEM
;  FILE REFERENCE: 2752-17
;  CURRENT APPLICATION NUMBER: US/09/790,812A
;  CURRENT FILING DATE:  2001-02-23
;  PRIOR APPLICATION NUMBER: 09/576,824
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 08/723,425
;  PRIOR FILING DATE: 1996-09-30
;  PRIOR APPLICATION NUMBER: 09/146,028
;  PRIOR FILING DATE: 1993-11-22
;  PRIOR APPLICATION NUMBER: PCT/EP93/00517
;  PRIOR FILING DATE: 1993-03-08
;  PRIOR APPLICATION NUMBER: EP 92400598.6
;  PRIOR FILING DATE: 1992-03-06
;  NUMBER OF SEQ ID NOS: 600
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 213
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Hepatitis C virus
US-09-790-812A-213


  Query Match            39.7%;  Score 23;  DB 5;  Length 9;
  Best Local Similarity   55.6%;  Pred. No. 2.9e+05;
  Matches    5;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 LMPFPGPLL 10
           |:|  || |
Db       1 LLPRRGPRL 9


Search completed: October 23, 2001, 13:36:58
Job time: 170 sec
```

```
;  PRIOR APPLICATION NUMBER: PCT/EP93/00517
;  PRIOR FILING DATE: 1993-03-08
;  PRIOR APPLICATION NUMBER: EP 92400598.6
;  PRIOR FILING DATE: 1992-03-06
;  NUMBER OF SEQ ID NOS: 600
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 213
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Hepatitis C virus
US-09-790-497A-213


  Query Match             39.7%;  Score 23;  DB 5;  Length 9;
  Best Local Similarity   55.6%;  Pred. No. 2.9e+05;
  Matches    5;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 LMPFPGPLL 10
           |:|  || |
Db       1 LLPRRGPRL 9


RESULT  10
US-09-790-497A-492
; Sequence 492, Application US/09790497A
; GENERAL INFORMATION:
;  APPLICANT: De Leys, Robert
;  TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES CORRESPONDING
;  TITLE OF INVENTION:  TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR USE IN
;  TITLE OF INVENTION:  A PROCESS FOR DETERMINATION OF ANTIBODIES OF
;  TITLE OF INVENTION:  BIOTINYLATED PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT
;  TITLE OF INVENTION:  EPITOPES, A PROCESS FOR PREPARING THEM AND COMPOSITIONS
;  TITLE OF INVENTION:  CONTAINING THEM
;  FILE REFERENCE: 2752-16
;  CURRENT APPLICATION NUMBER: US/09/790,497A
;  CURRENT FILING DATE:  2001-02-23
;  PRIOR APPLICATION NUMBER: 09/576,824
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 08/723,425
;  PRIOR FILING DATE: 1996-09-30
;  PRIOR APPLICATION NUMBER: 09/146,028
;  PRIOR FILING DATE: 1993-11-22
;  PRIOR APPLICATION NUMBER: PCT/EP93/00517
;  PRIOR FILING DATE: 1993-03-08
;  PRIOR APPLICATION NUMBER: EP 92400598.6
;  PRIOR FILING DATE: 1992-03-06
;  NUMBER OF SEQ ID NOS: 600
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 492
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Hepatitis C virus
US-09-790-497A-492


  Query Match             39.7%;  Score 23;  DB 5;  Length 9;
  Best Local Similarity   50.0%;  Pred. No. 2.9e+05;
  Matches    4;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 HLMPFPGP 8
           :|:|  ||
Db       2 YLLPRRGP 9

RESULT  11
US-09-790-497A-493
; Sequence 493, Application US/09790497A
; GENERAL INFORMATION:
;  APPLICANT: De Leys, Robert
;  TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES CORRESPONDING
;  TITLE OF INVENTION:  TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR USE IN
;  TITLE OF INVENTION:  A PROCESS FOR DETERMINATION OF ANTIBODIES OF
;  TITLE OF INVENTION:  BIOTINYLATED PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT
```

```
;  TITLE OF INVENTION:  EPITOPES, A PROCESS FOR PREPARING THEM AND COMPOSITIONS
;  TITLE OF INVENTION:  CONTAINING THEM
;  FILE REFERENCE: 2752-16
;  CURRENT APPLICATION NUMBER: US/09/790,497A
;  CURRENT FILING DATE:  2001-02-23
;  PRIOR APPLICATION NUMBER: 09/576,824
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 08/723,425
;  PRIOR FILING DATE: 1996-09-30
;  PRIOR APPLICATION NUMBER: 09/146,028
;  PRIOR FILING DATE: 1993-11-22
;  PRIOR APPLICATION NUMBER: PCT/EP93/00517
;  PRIOR FILING DATE: 1993-03-08
;  PRIOR APPLICATION NUMBER: EP 92400598.6
;  PRIOR FILING DATE: 1992-03-06
;  NUMBER OF SEQ ID NOS: 600
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 493
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Hepatitis C virus
US-09-790-497A-493


  Query Match             39.7%;  Score 23;  DB 5;  Length 9;
  Best Local Similarity   50.0%;  Pred. No. 2.9e+05;
  Matches    4;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 HLMPFPGP 8
           :|:|  ||
Db       1 YLLPRRGP 8


RESULT  12
US-09-790-497A-494
; Sequence 494, Application US/09790497A
; GENERAL INFORMATION:
;  APPLICANT: De Leys, Robert
;  TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES CORRESPONDING
;  TITLE OF INVENTION:  TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR USE IN
;  TITLE OF INVENTION:  A PROCESS FOR DETERMINATION OF ANTIBODIES OF
;  TITLE OF INVENTION:  BIOTINYLATED PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORT
;  TITLE OF INVENTION:  EPITOPES, A PROCESS FOR PREPARING THEM AND COMPOSITIONS
;  TITLE OF INVENTION:  CONTAINING THEM
;  FILE REFERENCE: 2752-16
;  CURRENT APPLICATION NUMBER: US/09/790,497A
;  CURRENT FILING DATE:  2001-02-23
;  PRIOR APPLICATION NUMBER: 09/576,824
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 08/723,425
;  PRIOR FILING DATE: 1996-09-30
;  PRIOR APPLICATION NUMBER: 09/146,028
;  PRIOR FILING DATE: 1993-11-22
;  PRIOR APPLICATION NUMBER: PCT/EP93/00517
;  PRIOR FILING DATE: 1993-03-08
;  PRIOR APPLICATION NUMBER: EP 92400598.6
;  PRIOR FILING DATE: 1992-03-06
;  NUMBER OF SEQ ID NOS: 600
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 494
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Hepatitis C virus
US-09-790-497A-494


  Query Match             39.7%;  Score 23;  DB 5;  Length 9;
  Best Local Similarity   55.6%;  Pred. No. 2.9e+05;
  Matches    5;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 LMPFPGPLL 10
           |:|  || |
Db       1 LLPRRGPRL 9
```

```
;  PRIOR FILING DATE: 1999-12-03
;  PRIOR APPLICATION NUMBER: US 60/124,095
;  PRIOR FILING DATE: 1999-03-12
;  PRIOR APPLICATION NUMBER: US 60/138,598
;  PRIOR FILING DATE: 1999-06-11
;  PRIOR APPLICATION NUMBER: US 60/168,665
;  PRIOR FILING DATE: 1999-12-03
;  PRIOR APPLICATION NUMBER: US 60/125,360
;  PRIOR FILING DATE: 1999-03-19
;  PRIOR APPLICATION NUMBER: US 60/138,626
;  PRIOR FILING DATE: 1999-06-11
;  PRIOR APPLICATION NUMBER: US 60/168,662
;  PRIOR FILING DATE: 1999-12-03
;  PRIOR APPLICATION NUMBER: US 60/124,144
;  PRIOR FILING DATE: 1999-03-12
;  PRIOR APPLICATION NUMBER: US 60/138,574
;  PRIOR FILING DATE: 1999-06-11
;  PRIOR APPLICATION NUMBER: US 60/168,667
;  PRIOR FILING DATE: 1999-12-03
;  PRIOR APPLICATION NUMBER: US 60/124,142
;  PRIOR FILING DATE: 1999-03-12
;  PRIOR APPLICATION NUMBER: US 60/138,597
;  PRIOR FILING DATE: 1999-06-11
;  PRIOR APPLICATION NUMBER: US 60/168,666
;  PRIOR FILING DATE: 1999-12-03
;  PRIOR APPLICATION NUMBER: US 60/125,359
;  PRIOR FILING DATE: 1999-03-19
;  PRIOR APPLICATION NUMBER: US 60/168,664
;  PRIOR FILING DATE: 1999-12-03
;  PRIOR APPLICATION NUMBER: US 60/126,051
;  PRIOR FILING DATE: 1999-03-23
;  PRIOR APPLICATION NUMBER: US 60/169,906

  Query Match             39.7%;  Score 23;  DB 5;  Length 6;
  Best Local Similarity   80.0%;  Pred. No. 2.9e+05;
  Matches    4;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      6 PGPLL 10
          ||| |
Db      2 PGPCL 6

RESULT   7
US-09-790-497A-211
; Sequence 211, Application US/09790497A
; GENERAL INFORMATION:
;  APPLICANT: De Leys, Robert
;  TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES CORRESPONDING
;  TITLE OF INVENTION:  TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR USE IN
;  TITLE OF INVENTION:  A PROCESS FOR DETERMINATION OF ANTIBODIES OF
;  TITLE OF INVENTION:  BIOTINYLATED PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT
;  TITLE OF INVENTION:  EPITOPES, A PROCESS FOR PREPARING THEM AND COMPOSITIONS
;  TITLE OF INVENTION:  CONTAINING THEM
;  FILE REFERENCE: 2752-16
;  CURRENT APPLICATION NUMBER: US/09/790,497A
;  CURRENT FILING DATE:  2001-02-23
;  PRIOR APPLICATION NUMBER: 09/576,824
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 08/723,425
;  PRIOR FILING DATE: 1996-09-30
;  PRIOR APPLICATION NUMBER: 09/146,028
;  PRIOR FILING DATE: 1993-11-22
;  PRIOR APPLICATION NUMBER: PCT/EP93/00517
;  PRIOR FILING DATE: 1993-03-08
;  PRIOR APPLICATION NUMBER: EP 92400598.6
;  PRIOR FILING DATE: 1992-03-06
;  NUMBER OF SEQ ID NOS: 600
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 211
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Hepatitis C virus
US-09-790-497A-211

  Query Match             39.7%;  Score 23;  DB 5;  Length 9;
  Best Local Similarity   50.0%;  Pred. No. 2.9e+05;
  Matches    4;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 HLMPFPGP 8
          :|:|  ||
Db      2 YLLPRRGP 9

RESULT   8
US-09-790-497A-212
; Sequence 212, Application US/09790497A
; GENERAL INFORMATION:
;  APPLICANT: De Leys, Robert
;  TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES CORRESPONDING
;  TITLE OF INVENTION:  TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR USE IN
;  TITLE OF INVENTION:  A PROCESS FOR DETERMINATION OF ANTIBODIES OF
;  TITLE OF INVENTION:  BIOTINYLATED PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORT
;  TITLE OF INVENTION:  EPITOPES, A PROCESS FOR PREPARING THEM AND COMPOSITIONS
;  TITLE OF INVENTION:  CONTAINING THEM
;  FILE REFERENCE: 2752-16
;  CURRENT APPLICATION NUMBER: US/09/790,497A
;  CURRENT FILING DATE:  2001-02-23
;  PRIOR APPLICATION NUMBER: 09/576,824
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 08/723,425
;  PRIOR FILING DATE: 1996-09-30
;  PRIOR APPLICATION NUMBER: 09/146,028
;  PRIOR FILING DATE: 1993-11-22
;  PRIOR APPLICATION NUMBER: PCT/EP93/00517
;  PRIOR FILING DATE: 1993-03-08
;  PRIOR APPLICATION NUMBER: EP 92400598.6
;  PRIOR FILING DATE: 1992-03-06
;  NUMBER OF SEQ ID NOS: 600
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 212
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Hepatitis C virus
US-09-790-497A-212

  Query Match             39.7%;  Score 23;  DB 5;  Length 9;
  Best Local Similarity   50.0%;  Pred. No. 2.9e+05;
  Matches    4;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 HLMPFPGP 8
          :|:|  ||
Db      1 YLLPRRGP 8

RESULT   9
US-09-790-497A-213
; Sequence 213, Application US/09790497A
; GENERAL INFORMATION:
;  APPLICANT: De Leys, Robert
;  TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES CORRESPONDING
;  TITLE OF INVENTION:  TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR USE IN
;  TITLE OF INVENTION:  A PROCESS FOR DETERMINATION OF ANTIBODIES OF
;  TITLE OF INVENTION:  BIOTINYLATED PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORT
;  TITLE OF INVENTION:  EPITOPES, A PROCESS FOR PREPARING THEM AND COMPOSITIONS
;  TITLE OF INVENTION:  CONTAINING THEM
;  FILE REFERENCE: 2752-16
;  CURRENT APPLICATION NUMBER: US/09/790,497A
;  CURRENT FILING DATE:  2001-02-23
;  PRIOR APPLICATION NUMBER: 09/576,824
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 08/723,425
;  PRIOR FILING DATE: 1996-09-30
;  PRIOR APPLICATION NUMBER: 09/146,028
;  PRIOR FILING DATE: 1993-11-22
```

```
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06042
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06014
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06013
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06049
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06057
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06824
;  PRIOR FILING DATE: 2000-03-16
;  PRIOR APPLICATION NUMBER: PCT/US00/06765
;  PRIOR FILING DATE: 2000-03-16
;  PRIOR APPLICATION NUMBER: PCT/US00/06792
;  PRIOR FILING DATE: 2000-03-16
;  PRIOR APPLICATION NUMBER: PCT/US00/06830
;  PRIOR FILING DATE: 2000-03-16
;  PRIOR APPLICATION NUMBER: PCT/US00/06782
;  PRIOR FILING DATE: 2000-03-16
;  PRIOR APPLICATION NUMBER: PCT/US00/06822
;  PRIOR FILING DATE: 2000-03-16
;  PRIOR APPLICATION NUMBER: PCT/US00/06791
;  PRIOR FILING DATE: 2000-03-16
;  PRIOR APPLICATION NUMBER: PCT/US00/06828
;  PRIOR FILING DATE: 2000-03-16
;  PRIOR APPLICATION NUMBER: PCT/US00/06823
;  PRIOR FILING DATE: 2000-03-16
;  PRIOR APPLICATION NUMBER: PCT/US00/06781
;  PRIOR FILING DATE: 2000-03-16
;  PRIOR APPLICATION NUMBER: PCT/US00/07505
;  PRIOR FILING DATE: 2000-03-22
;  PRIOR APPLICATION NUMBER: PCT/US00/07440
;  PRIOR FILING DATE: 2000-03-22
;  PRIOR APPLICATION NUMBER: PCT/US00/07506
;  PRIOR FILING DATE: 2000-03-22
;  PRIOR APPLICATION NUMBER: PCT/US00/07507
;  PRIOR FILING DATE: 2000-03-22
;  PRIOR APPLICATION NUMBER: PCT/US00/07535
;  PRIOR FILING DATE: 2000-03-22
;  PRIOR APPLICATION NUMBER: PCT/US00/07525
;  PRIOR FILING DATE: 2000-03-22
;  PRIOR APPLICATION NUMBER: PCT/US00/07534
;  PRIOR FILING DATE: 2000-03-22
;  PRIOR APPLICATION NUMBER: PCT/US00/07483
;  PRIOR FILING DATE: 2000-03-22
;  PRIOR APPLICATION NUMBER: PCT/US00/07526
;  PRIOR FILING DATE: 2000-03-22
;  PRIOR APPLICATION NUMBER: PCT/US00/07527
;  PRIOR FILING DATE: 2000-03-22
;  PRIOR APPLICATION NUMBER: PCT/US00/07661
;  PRIOR FILING DATE: 2000-03-23
;  PRIOR APPLICATION NUMBER: PCT/US00/07579
;  PRIOR FILING DATE: 2000-03-23
;  PRIOR APPLICATION NUMBER: PCT/US00/07723
;  PRIOR FILING DATE: 2000-03-23
;  PRIOR APPLICATION NUMBER: PCT/US00/07724
;  PRIOR FILING DATE: 2000-03-23
;  PRIOR APPLICATION NUMBER: PCT/US00/14929
;  PRIOR FILING DATE: 2000-06-01
;  PRIOR APPLICATION NUMBER: PCT/US00/07722
;  PRIOR FILING DATE: 2000-03-23
;  PRIOR APPLICATION NUMBER: PCT/US00/07578
;  PRIOR FILING DATE: 2000-03-23
;  PRIOR APPLICATION NUMBER: PCT/US00/07726
;  PRIOR FILING DATE: 2000-03-23
;  PRIOR APPLICATION NUMBER: PCT/US00/07677
;  PRIOR FILING DATE: 2000-03-23
;  PRIOR APPLICATION NUMBER: PCT/US00/07725
;  PRIOR FILING DATE: 2000-03-23
;  PRIOR APPLICATION NUMBER: PCT/US00/09070
;  PRIOR FILING DATE: 2000-04-06
;  PRIOR APPLICATION NUMBER: PCT/US00/08982
;  PRIOR FILING DATE: 2000-04-06
;  PRIOR APPLICATION NUMBER: PCT/US00/08983
;  PRIOR FILING DATE: 2000-04-06
;  PRIOR APPLICATION NUMBER: PCT/US00/09067
;  PRIOR FILING DATE: 2000-04-06
;  PRIOR APPLICATION NUMBER: PCT/US00/09066
;  PRIOR FILING DATE: 2000-04-06
;  PRIOR APPLICATION NUMBER: PCT/US00/09068
;  PRIOR FILING DATE: 2000-04-06
;  PRIOR APPLICATION NUMBER: PCT/US00/08981
;  PRIOR FILING DATE: 2000-04-06
;  PRIOR APPLICATION NUMBER: PCT/US00/08980
;  PRIOR FILING DATE: 2000-04-06
;  PRIOR APPLICATION NUMBER: PCT/US00/09071
;  PRIOR FILING DATE: 2000-04-06
;  PRIOR APPLICATION NUMBER: PCT/US00/09069
;  PRIOR FILING DATE: 2000-04-06
;  PRIOR APPLICATION NUMBER: PCT/US00/15136
;  PRIOR FILING DATE: 2000-06-01
;  PRIOR APPLICATION NUMBER: PCT/US00/14926
;  PRIOR FILING DATE: 2000-06-01
;  PRIOR APPLICATION NUMBER: PCT/US00/14963
;  PRIOR FILING DATE: 2000-06-01
;  PRIOR APPLICATION NUMBER: PCT/US00/15135
;  PRIOR FILING DATE: 2000-06-01
;  PRIOR APPLICATION NUMBER: PCT/US00/14934
;  PRIOR FILING DATE: 2000-06-01
;  PRIOR APPLICATION NUMBER: PCT/US00/14933
;  PRIOR FILING DATE: 2000-06-01
;  PRIOR APPLICATION NUMBER: PCT/US00/15137
;  PRIOR FILING DATE: 2000-06-01
;  PRIOR APPLICATION NUMBER: PCT/US00/14928
;  PRIOR FILING DATE: 2000-06-01
;  PRIOR APPLICATION NUMBER: PCT/US00/14973
;  PRIOR FILING DATE: 2000-06-01
;  PRIOR APPLICATION NUMBER: PCT/US00/14964
;  PRIOR FILING DATE: 2000-06-01
;  PRIOR APPLICATION NUMBER: PCT/US00/26376
;  PRIOR FILING DATE: 2000-09-26
;  PRIOR APPLICATION NUMBER: PCT/US00/26371
;  PRIOR FILING DATE: 2000-09-26
;  PRIOR APPLICATION NUMBER: PCT/US00/26324
;  PRIOR FILING DATE: 2000-09-26
;  PRIOR APPLICATION NUMBER: PCT/US00/26323
;  PRIOR FILING DATE: 2000-09-26
;  PRIOR APPLICATION NUMBER: PCT/US00/26337
;  PRIOR FILING DATE: 2000-09-26
;  PRIOR APPLICATION NUMBER: PCT/US01/13318
;  PRIOR FILING DATE: 2001-04-27
;  PRIOR APPLICATION NUMBER: US 60/124,146
;  PRIOR FILING DATE: 1999-03-12
;  PRIOR APPLICATION NUMBER: US 60/167,061
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/124,093
;  PRIOR FILING DATE: 1999-03-12
;  PRIOR APPLICATION NUMBER: US 60/166,989
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/124,145
;  PRIOR FILING DATE: 1999-03-12
;  PRIOR APPLICATION NUMBER: US 60/168,654
;  PRIOR FILING DATE: 1999-12-03
;  PRIOR APPLICATION NUMBER: US 60/124,099
;  PRIOR FILING DATE: 1999-03-12
;  PRIOR APPLICATION NUMBER: US 60/168,661
;  PRIOR FILING DATE: 1999-12-03
;  PRIOR APPLICATION NUMBER: US 60/124,096
;  PRIOR FILING DATE: 1999-03-12
;  PRIOR APPLICATION NUMBER: US 60/168,622
;  PRIOR FILING DATE: 1999-12-03
;  PRIOR APPLICATION NUMBER: US 60/124,143
;  PRIOR FILING DATE: 1999-03-12
;  PRIOR APPLICATION NUMBER: US 60/168,663
```

```
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-719-554-109


  Query Match             50.0%;  Score 29;  DB 5;  Length 10;
  Best Local Similarity   71.4%;  Pred. No. 2e+02;
  Matches    5;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          3 MPFPGPL 9
              :|| |||
Db          1 LPFLGPL 7


RESULT   3
PCT-US01-22458-10
; Sequence 10, Application PC/TUS0122458
; GENERAL INFORMATION:
;  APPLICANT: Merck & Co., Inc.
;  TITLE OF INVENTION: DOG MELANIN-CONCENTRATING HORMONE
;  TITLE OF INVENTION:  RECEPTOR
;  FILE REFERENCE: 20668 PCT
;  CURRENT APPLICATION NUMBER: PCT/US01/22458
;  CURRENT FILING DATE:  2001-07-17
;  PRIOR APPLICATION NUMBER: 60/219,669
;  PRIOR FILING DATE: 2000-07-21
;  NUMBER OF SEQ ID NOS: 18
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 10
;   LENGTH: 10
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Corresponds to a dog MCH receptor region
PCT-US01-22458-10


  Query Match             44.8%;  Score 26;  DB 1;  Length 10;
  Best Local Similarity   100.0%;  Pred. No. 5.7e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          4 PFPG 7
              ||||
Db          1 PFPG 4


RESULT   4
US-09-832-723-66
; Sequence 66, Application US/09832723
; GENERAL INFORMATION:
;  APPLICANT: Estell, David A.
;  APPLICANT:  Chen, Yiyou
;  APPLICANT:  Murray, Christopher J.
;  APPLICANT:  Tijerina, Pilar
;  TITLE OF INVENTION: METHODS FOR SELECTIVE TARGETING
;  FILE REFERENCE: GC617-2
;  CURRENT APPLICATION NUMBER: US/09/832,723
;  CURRENT FILING DATE:  2001-04-11
;  PRIOR APPLICATION NUMBER: US 60/197,259
;  PRIOR FILING DATE: 2000-04-14
;  NUMBER OF SEQ ID NOS: 117
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 66
;   LENGTH: 7
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: peptides screened from a phage display random
;   OTHER INFORMATION: peptide library
US-09-832-723-66


  Query Match             41.4%;  Score 24;  DB 5;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 2.9e+05;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          6 PGPL 9
              ||||
Db          2 PGPL 5


RESULT   5
US-09-832-723-70
; Sequence 70, Application US/09832723
; GENERAL INFORMATION:
;  APPLICANT: Estell, David A.
;  APPLICANT:  Chen, Yiyou
;  APPLICANT:  Murray, Christopher J.
;  APPLICANT:  Tijerina, Pilar
;  TITLE OF INVENTION: METHODS FOR SELECTIVE TARGETING
;  FILE REFERENCE: GC617-2
;  CURRENT APPLICATION NUMBER: US/09/832,723
;  CURRENT FILING DATE:  2001-04-11
;  PRIOR APPLICATION NUMBER: US 60/197,259
;  PRIOR FILING DATE: 2000-04-14
;  NUMBER OF SEQ ID NOS: 117
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 70
;   LENGTH: 7
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: peptides screened from a phage display random
;   OTHER INFORMATION: peptide library
;   NAME/KEY: VARIANT
;   LOCATION: (1)...(7)
;   OTHER INFORMATION: Xaa = Any Amino Acid
;   NAME/KEY: VARIANT
;   LOCATION: (1)...(7)
;   OTHER INFORMATION: Xaa = Any Amino Acid
US-09-832-723-70


  Query Match             41.4%;  Score 24;  DB 5;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 2.9e+05;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          6 PGPL 9
              ||||
Db          2 PGPL 5


RESULT   6
US-09-950-083-4410
; Sequence 4410, Application US/09950083
; GENERAL INFORMATION:
;  APPLICANT: Rosen, et. al
;  TITLE OF INVENTION: Human Secreted Proteins
;  FILE REFERENCE: PS805
;  CURRENT APPLICATION NUMBER: US/09/950,083
;  CURRENT FILING DATE:  2001-09-12
;  PRIOR APPLICATION NUMBER: 60/278,650
;  PRIOR FILING DATE: 2001-03-27
;  PRIOR APPLICATION NUMBER: 09/833,245
;  PRIOR FILING DATE: 2001-04-12
;  PRIOR APPLICATION NUMBER: PCT/US01/11988
;  PRIOR FILING DATE: 2001-04-12
;  PRIOR APPLICATION NUMBER: PCT/US00/06043
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06012
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06058
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06044
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06059
```

```
                         GenCore version 4.5
                Copyright (c) 1993 - 2000  Compugen Ltd.

OM protein - protein search, using sw model

Run on:          October 23, 2001, 13:34:08 ; Search time 15.45 Seconds
                                              (without alignments)
                                              33.351 Million cell updates/sec

Title:           US-09-625-963-3
Perfect score:   58
Sequence:        1 HLMPFPGPLL 10

Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched:        339663 seqs, 51527020 residues

Total number of hits satisfying chosen parameters:      64260

Minimum DB seq length: 0
Maximum DB seq length: 10

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       Pending_Patents_AA_New:*
                1:  /cgn2_6/ptodata/2/paa/PCT_NEW_COMB.pep:*
                2:  /cgn2_6/ptodata/2/paa/US06_NEW_COMB.pep:*
                3:  /cgn2_6/ptodata/2/paa/US07_NEW_COMB.pep:*
                4:  /cgn2_6/ptodata/2/paa/US08_NEW_COMB.pep:*
                5:  /cgn2_6/ptodata/2/paa/US09_NEW_COMB.pep:*
                6:  /cgn2_6/ptodata/2/paa/US60_NEW_COMB.pep:*

     Pred. No. is the number of results predicted by chance to have a
     score greater than or equal to the score of the result being printed,
     and is derived by analysis of the total score distribution.

                              SUMMARIES
                %
Result        Query
  No.  Score  Match Length DB  ID                         Description
-------------------------------------------------------------------------
    1     29   50.0     10  5  US-09-719-554-91            Sequence 91, Appl
    2     29   50.0     10  5  US-09-719-554-109           Sequence 109, App
    3     26   44.8     10  1  PCT-US01-22458-10           Sequence 10, Appl
    4     24   41.4      7  5  US-09-832-723-66            Sequence 66, Appl
    5     24   41.4      7  5  US-09-832-723-70            Sequence 70, Appl
    6     23   39.7      6  5  US-09-950-083-4410          Sequence 4410, Ap
    7     23   39.7      9  5  US-09-790-497A-211          Sequence 211, App
    8     23   39.7      9  5  US-09-790-497A-212          Sequence 212, App
    9     23   39.7      9  5  US-09-790-497A-213          Sequence 213, App
   10     23   39.7      9  5  US-09-790-497A-492          Sequence 492, App
   11     23   39.7      9  5  US-09-790-497A-493          Sequence 493, App
   12     23   39.7      9  5  US-09-790-497A-494          Sequence 494, App
   13     23   39.7      9  5  US-09-790-812A-211          Sequence 211, App
   14     23   39.7      9  5  US-09-790-812A-212          Sequence 212, App
   15     23   39.7      9  5  US-09-790-812A-213          Sequence 213, App
   16     23   39.7      9  5  US-09-790-812A-492          Sequence 492, App
   17     23   39.7      9  5  US-09-790-812A-493          Sequence 493, App
   18     23   39.7      9  5  US-09-790-812A-494          Sequence 494, App
   19   22.5   38.8     10  1  PCT-US01-22458-7            Sequence 7, Appli
   20     22   37.9      8  5  US-09-633-364-91            Sequence 91, Appl
   21     22   37.9      8  5  US-09-633-364-92            Sequence 92, Appl
   22     22   37.9      8  5  US-09-633-364-2107          Sequence 2107, Ap
   23     22   37.9      8  5  US-09-633-364-2110          Sequence 2110, Ap
   24     22   37.9      8  5  US-09-633-364-2497          Sequence 2497, Ap
   25     22   37.9      8  5  US-09-633-364-2498          Sequence 2498, Ap
   26     22   37.9      8  5  US-09-633-364-2747          Sequence 2747, Ap
   27     22   37.9      8  5  US-09-633-364-2748          Sequence 2748, Ap
   28     22   37.9      8  5  US-09-633-364-3265          Sequence 3265, Ap
   29     22   37.9      8  5  US-09-633-364-3268          Sequence 3268, Ap
   30     22   37.9      8  5  US-09-633-364-3987          Sequence 3987, Ap
   31     22   37.9      8  5  US-09-633-364-3988          Sequence 3988, Ap
   32     22   37.9      8  5  US-09-633-364-4418          Sequence 4418, Ap
   33     22   37.9      8  5  US-09-633-364-4419          Sequence 4419, Ap
   34     22   37.9      8  5  US-09-633-364-5378          Sequence 5378, Ap
   35     22   37.9      8  5  US-09-633-364-5379          Sequence 5379, Ap
   36     22   37.9      8  5  US-09-832-355A-10           Sequence 10, Appl
   37     22   37.9      9  1  PCT-US00-08879A-24          Sequence 24, Appl
   38     22   37.9      9  5  US-09-633-364-1876          Sequence 1876, Ap
   39     22   37.9      9  5  US-09-633-364-1878          Sequence 1878, Ap
   40     22   37.9      9  5  US-09-633-364-2499          Sequence 2499, Ap
   41     22   37.9      9  5  US-09-633-364-2500          Sequence 2500, Ap
   42     22   37.9      9  5  US-09-633-364-3007          Sequence 3007, Ap
   43     22   37.9      9  5  US-09-633-364-3008          Sequence 3008, Ap
   44     22   37.9      9  5  US-09-633-364-4049          Sequence 4049, Ap
   45     22   37.9      9  5  US-09-633-364-6544          Sequence 6544, Ap

                              ALIGNMENTS

RESULT   1
US-09-719-554-91
; Sequence 91, Application US/09719554
; GENERAL INFORMATION:
;  APPLICANT: ALLIEL, Patrick
;  APPLICANT:  PERIN, Jean-Pierre
;  APPLICANT:  RIEGER, Francois
;  TITLE OF INVENTION: NUCLEIC ACID SEQUENCE AND DEDUCED PROTEIN SEQUENCE FAMILY WITH
;  TITLE OF INVENTION:  ENDOGENOUS RETROVIRAL MOTIFS
;  FILE REFERENCE: 200936US0PCT
;  CURRENT APPLICATION NUMBER: US/09/719,554
;  CURRENT FILING DATE:  2001-12-26
;  PRIOR APPLICATION NUMBER: PCT/FR99/01513
;  PRIOR FILING DATE: 1999-06-25
;  NUMBER OF SEQ ID NOS: 122
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 91
;   LENGTH: 10
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-719-554-91

  Query Match             50.0%;  Score 29;  DB 5;  Length 10;
  Best Local Similarity   71.4%;  Pred. No. 2e+02;
  Matches    5;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       3 MPFPGPL 9
           :|| |||
Db       1 LPFLGPL 7

RESULT   2
US-09-719-554-109
; Sequence 109, Application US/09719554
; GENERAL INFORMATION:
;  APPLICANT: ALLIEL, Patrick
;  APPLICANT:  PERIN, Jean-Pierre
;  APPLICANT:  RIEGER, Francois
;  TITLE OF INVENTION: NUCLEIC ACID SEQUENCE AND DEDUCED PROTEIN SEQUENCE FAMILY WITH
;  TITLE OF INVENTION:  ENDOGENOUS RETROVIRAL MOTIFS
;  FILE REFERENCE: 200936US0PCT
;  CURRENT APPLICATION NUMBER: US/09/719,554
;  CURRENT FILING DATE:  2001-12-26
;  PRIOR APPLICATION NUMBER: PCT/FR99/01513
;  PRIOR FILING DATE: 1999-06-25
;  NUMBER OF SEQ ID NOS: 122
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 109
;   LENGTH: 10
```

```
RESULT  15
US-08-073-205-196
; Sequence 196, Application US/08073205
;  GENERAL INFORMATION:
;    APPLICANT:  GREY, Howard M.
;    APPLICANT:  SETTE, Alessandro
;    APPLICANT:  SIDNEY, John
;    APPLICANT:  KAST, Wijbe M.
;    TITLE OF INVENTION:  HLA-A2.1 BINDING PEPTIDES AND THEIR USES
;    NUMBER OF SEQUENCES:  696
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend Khourie and Crew
;      STREET:  Steuart Street Tower, One Market Plaza
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  US
;      ZIP:  94105-1493
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/073,205
;      FILING DATE:  04-JUN-1993
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/027,146
;      FILING DATE:  05-MAR-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Bastian, Kevin L.
;      REGISTRATION NUMBER:  34,774
;      REFERENCE/DOCKET NUMBER:  14137-58-1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 543-9600
;      TELEFAX:  (415) 543-5043
;  INFORMATION FOR SEQ ID NO:  196:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-073-205-196


  Query Match             53.4%;  Score 31;  DB 4;  Length 9;
  Best Local Similarity   83.3%;  Pred. No. 2.6e+06;
  Matches    5;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          4 PFPGPL 9
              | ||||
Db          1 PLPGPL 6


Search completed: October 23, 2001, 13:36:36
Job time: 168 sec
```

```
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/461,788
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/224,917
;      FILING DATE:  08-APR-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Peries, Rohan
;      REGISTRATION NUMBER:  35,752
;      REFERENCE/DOCKET NUMBER:  28060
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415)-852-1698
;      TELEFAX:  (415)-496-3529
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    HYPOTHETICAL:  NO
;    FRAGMENT TYPE:  internal
US-08-461-788-8


  Query Match            56.9%;  Score 33;  DB 8;  Length 9;
  Best Local Similarity  100.0%;  Pred. No. 2.6e+06;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       4 PFPGP 8
           |||||
Db       1 PFPGP 5


RESULT 13
US-08-464-505-8
; Sequence 8, Application US/08464505
;  GENERAL INFORMATION:
;    APPLICANT:  Lou, Lillian Lien-Li
;    APPLICANT:  Barnett, Jimmy Wayne
;    TITLE OF INVENTION:  Cloning and Expression of Human GMP
;    TITLE OF INVENTION:  Synthetase, its use in Screening for Inhibitors of
;    TITLE OF INVENTION:  Human
;    TITLE OF INVENTION:  GMP Synthetase and Inhibitors of Human GMP Synthetase
;    NUMBER OF SEQUENCES:  11
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Syntex (USA) Inc.
;      STREET:  3401 Hillview Avenue
;      CITY:  Palo Alto
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/464,505
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/224,917
;      FILING DATE:  08-APR-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Peries, Rohan
;      REGISTRATION NUMBER:  35,752
;      REFERENCE/DOCKET NUMBER:  28060
```

```
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415)-852-1698
;      TELEFAX:  (415)-496-3529
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    HYPOTHETICAL:  NO
;    FRAGMENT TYPE:  internal
US-08-464-505-8


  Query Match            56.9%;  Score 33;  DB 8;  Length 9;
  Best Local Similarity  100.0%;  Pred. No. 2.6e+06;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       4 PFPGP 8
           |||||
Db       1 PFPGP 5


RESULT 14
US-08-027-146-196
; Sequence 196, Application US/08027146
;  GENERAL INFORMATION:
;    APPLICANT:  GRAY, Howard M.
;    APPLICANT:  SETTE, Alessandro
;    APPLICANT:  SIDNEY, John
;    APPLICANT:  KAST, W. M.
;    TITLE OF INVENTION:  HLA-A2.1 BINDING PEPTIDES AND THEIR USES
;    NUMBER OF SEQUENCES:  390
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend Khourie and Crew
;      STREET:  Steuart Street Tower, One Market Plaza
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  US
;      ZIP:  94105-1493
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/027,146
;      FILING DATE:  19930305
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Bastian, Kevin L.
;      REGISTRATION NUMBER:  34,774
;      REFERENCE/DOCKET NUMBER:  14137-58
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 543-9600
;      TELEFAX:  (415) 543-5043
;  INFORMATION FOR SEQ ID NO:  196:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-027-146-196


  Query Match            53.4%;  Score 31;  DB 4;  Length 9;
  Best Local Similarity  83.3%;  Pred. No. 2.6e+06;
  Matches    5;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       4 PFPGPL 9
           | ||||
Db       1 PLPGPL 6
```

```
  Query Match             56.9%;  Score 33;  DB 17;  Length 6;
  Best Local Similarity   100.0%;  Pred. No. 2.6e+06;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       4 PFPGP 8
           |||||
Db       2 PFPGP 6


RESULT  10
US-08-461-462-8
; Sequence 8, Application US/08461462
;  GENERAL INFORMATION:
;    APPLICANT:  Lou, Lillian Lien-Li
;    APPLICANT:  Barnett, Jimmy Wayne
;    TITLE OF INVENTION:  Cloning and Expression of Human GMP
;    TITLE OF INVENTION:  Synthetase, its use in Screening for Inhibitors of
;    TITLE OF INVENTION:  Human
;    TITLE OF INVENTION:  GMP Synthetase and Inhibitors of Human GMP Synthetase
;    NUMBER OF SEQUENCES:  11
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Syntex (USA) Inc.
;      STREET:  3401 Hillview Avenue
;      CITY:  Palo Alto
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/461,462
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/224,917
;      FILING DATE:  08-APR-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Peries, Rohan
;      REGISTRATION NUMBER:  35,752
;      REFERENCE/DOCKET NUMBER:  28060
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415)-852-1698
;      TELEFAX:  (415)-496-3529
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    HYPOTHETICAL:  NO
;    FRAGMENT TYPE:  internal
US-08-461-462-8


  Query Match             56.9%;  Score 33;  DB 8;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.6e+06;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       4 PFPGP 8
           |||||
Db       1 PFPGP 5


RESULT  11
US-08-461-489-8
; Sequence 8, Application US/08461489
;  GENERAL INFORMATION:
;    APPLICANT:  Lou, Lillian Lien-Li
;    APPLICANT:  Barnett, Jimmy Wayne
;    TITLE OF INVENTION:  Cloning and Expression of Human GMP
;    TITLE OF INVENTION:  Synthetase, its use in Screening for Inhibitors of
;    TITLE OF INVENTION:  Human
;    TITLE OF INVENTION:  GMP Synthetase and Inhibitors of Human GMP Synthetase
;    NUMBER OF SEQUENCES:  11
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Syntex (USA) Inc.
;      STREET:  3401 Hillview Avenue
;      CITY:  Palo Alto
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/461,489
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/224,917
;      FILING DATE:  08-APR-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Peries, Rohan
;      REGISTRATION NUMBER:  35,752
;      REFERENCE/DOCKET NUMBER:  28060
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415)-852-1698
;      TELEFAX:  (415)-496-3529
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    HYPOTHETICAL:  NO
;    FRAGMENT TYPE:  internal
US-08-461-489-8


  Query Match             56.9%;  Score 33;  DB 8;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.6e+06;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       4 PFPGP 8
           |||||
Db       1 PFPGP 5


RESULT  12
US-08-461-788-8
; Sequence 8, Application US/08461788
;  GENERAL INFORMATION:
;    APPLICANT:  Lou, Lillian Lien-Li
;    APPLICANT:  Barnett, Jimmy Wayne
;    TITLE OF INVENTION:  Cloning and Expression of Human GMP
;    TITLE OF INVENTION:  Synthetase, its use in Screening for Inhibitors of
;    TITLE OF INVENTION:  Human
;    TITLE OF INVENTION:  GMP Synthetase and Inhibitors of Human GMP Synthetase
;    NUMBER OF SEQUENCES:  11
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Syntex (USA) Inc.
;      STREET:  3401 Hillview Avenue
;      CITY:  Palo Alto
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
```

```
;      NAME:  Cooper, Iver P.
;      REGISTRATION NUMBER:  28,005
;      REFERENCE/DOCKET NUMBER:  ANATHAN1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-628-5197
;      TELEFAX:  202-737-3528
;      TELEX:  248633
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  7 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-247-451-6

  Query Match             60.3%;  Score 35;  DB 6;  Length 7;
  Best Local Similarity   83.3%;  Pred. No. 2.6e+06;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      4 PFPGPL 9
          |||||:
Db      2 PFPGPI 7

RESULT   6
US-09-317-702-1
; Sequence 1, Application US/09317702A
; GENERAL INFORMATION:
;  APPLICANT: Ray, Michelle
;  APPLICANT: Venturini, Al
;  APPLICANT: Daiss, Jack
;  APPLICANT: Hinchman, Carolyn
;  TITLE OF INVENTION: DIAGNOSTIC MARKERS FOR HUMAN DISORDERS
;  FILE REFERENCE: CDS195 Jim Harrington
;  CURRENT APPLICATION NUMBER: US/09/317,702A
;  CURRENT FILING DATE: 1999-05-24
;  NUMBER OF SEQ ID NOS: 43
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 7
;   TYPE: PRT
;   ORGANISM: HUMAN
US-09-317-702-1

  Query Match             60.3%;  Score 35;  DB 17;  Length 7;
  Best Local Similarity   83.3%;  Pred. No. 2.6e+06;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      4 PFPGPL 9
          |||||:
Db      2 PFPGPI 7

RESULT   7
US-09-428-692-5
; Sequence 5, Application US/09428692
; GENERAL INFORMATION:
;  APPLICANT: Carr et al.
;  TITLE OF INVENTION: NOVEL CHIMERIC ANALGESIC PEPTIDES
;  FILE REFERENCE: 18475-016
;  CURRENT APPLICATION NUMBER: US/09/428,692
;  CURRENT FILING DATE:  1999-10-28
;  NUMBER OF SEQ ID NOS: 43
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
;   LENGTH: 7
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: recombinant
US-09-428-692-5
```

```
  Query Match             60.3%;  Score 35;  DB 18;  Length 7;
  Best Local Similarity   83.3%;  Pred. No. 2.6e+06;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      4 PFPGPL 9
          |||||:
Db      2 PFPGPI 7

RESULT   8
US-09-657-276-748
; Sequence 748, Application US/09657276
; GENERAL INFORMATION:
;  APPLICANT:  Conjuchem, Inc.
;  APPLICANT:  Bridon, Dominique
;  APPLICANT:  Ezrin, Alan
;  APPLICANT:  Milner, Peter
;  APPLICANT:  Holmes, Darren
;  APPLICANT:  Thibaudeau, Karen
;  TITLE OF INVENTION:  PROTECTION OF ENDOGENOUS THERAPEUTIC PEPTIDES FROM
;  TITLE OF INVENTION:  PEPTIDASE ACTIVITY THROUGH CONJUGATION TO BLOOD
;  TITLE OF INVENTION:  COMPONENTS
;  FILE REFERENCE:  2110
;  CURRENT APPLICATION NUMBER: US/09/657,276
;  CURRENT FILING DATE:  2000-09-07
;  PRIOR APPLICATION NUMBER:  60/134,406
;  PRIOR FILING DATE:  1999-05-17
;  PRIOR APPLICATION NUMBER:  60/153,406
;  PRIOR FILING DATE:  1999-09-10
;  PRIOR APPLICATION NUMBER:  60/159,783
;  PRIOR FILING DATE:  1999-10-18
;  NUMBER OF SEQ ID NOS:  1617
;  SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 748
;   LENGTH: 7
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Synthetic
;   OTHER INFORMATION: Peptide
US-09-657-276-748

  Query Match             60.3%;  Score 35;  DB 20;  Length 7;
  Best Local Similarity   83.3%;  Pred. No. 2.6e+06;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      4 PFPGPL 9
          |||||:
Db      2 PFPGPI 7

RESULT   9
US-09-317-702-2
; Sequence 2, Application US/09317702A
; GENERAL INFORMATION:
;  APPLICANT: Ray, Michelle
;  APPLICANT: Venturini, Al
;  APPLICANT: Daiss, Jack
;  APPLICANT: Hinchman, Carolyn
;  TITLE OF INVENTION: DIAGNOSTIC MARKERS FOR HUMAN DISORDERS
;  FILE REFERENCE: CDS195 Jim Harrington
;  CURRENT APPLICATION NUMBER: US/09/317,702A
;  CURRENT FILING DATE: 1999-05-24
;  NUMBER OF SEQ ID NOS: 43
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
;   LENGTH: 6
;   TYPE: PRT
;   ORGANISM: HUMAN
US-09-317-702-2
```

```
RESULT   2
US-09-317-702-6
; Sequence 6, Application US/09317702A
; GENERAL INFORMATION:
;  APPLICANT: Ray, Michelle
;  APPLICANT: Venturini, Al
;  APPLICANT: Daiss, Jack
;  APPLICANT: Hinchman, Carolyn
;  TITLE OF INVENTION: DIAGNOSTIC MARKERS FOR HUMAN DISORDERS
;  FILE REFERENCE: CDS195 Jim Harrington
;  CURRENT APPLICATION NUMBER: US/09/317,702A
;  CURRENT FILING DATE: 1999-05-24
;  NUMBER OF SEQ ID NOS: 43
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
;   LENGTH: 6
;   TYPE: PRT
;   ORGANISM: HUMAN
US-09-317-702-6


  Query Match             60.3%;  Score 35;  DB 17;  Length 6;
  Best Local Similarity   83.3%;  Pred. No. 2.6e+06;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       4 PFPGPL 9
           |||||:
Db       1 PFPGPI 6

RESULT   3
US-09-657-276-766
; Sequence 766, Application US/09657276
; GENERAL INFORMATION:
;  APPLICANT:  Conjuchem, Inc.
;  APPLICANT:  Bridon, Dominique
;  APPLICANT:  Ezrin, Alan
;  APPLICANT:  Milner, Peter
;  APPLICANT:  Holmes, Darren
;  APPLICANT:  Thibaudeau, Karen
;  TITLE OF INVENTION:  PROTECTION OF ENDOGENOUS THERAPEUTIC PEPTIDES FROM
;  TITLE OF INVENTION:  PEPTIDASE ACTIVITY THROUGH CONJUGATION TO BLOOD
;  TITLE OF INVENTION:  COMPONENTS
;  FILE REFERENCE:  2110
;  CURRENT APPLICATION NUMBER: US/09/657,276
;  CURRENT FILING DATE:  2000-09-07
;  PRIOR APPLICATION NUMBER:  60/134,406
;  PRIOR FILING DATE:  1999-05-17
;  PRIOR APPLICATION NUMBER:  60/153,406
;  PRIOR FILING DATE:  1999-09-10
;  PRIOR APPLICATION NUMBER:  60/159,783
;  PRIOR FILING DATE:  1999-10-18
;  NUMBER OF SEQ ID NOS:  1617
;  SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 766
;   LENGTH: 6
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Synthetic
;   OTHER INFORMATION: Peptide
US-09-657-276-766


  Query Match             60.3%;  Score 35;  DB 20;  Length 6;
  Best Local Similarity   83.3%;  Pred. No. 2.6e+06;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       4 PFPGPL 9
           |||||:
Db       1 PFPGPI 6
```

```
RESULT   4
US-09-657-276-767
; Sequence 767, Application US/09657276
; GENERAL INFORMATION:
;  APPLICANT:  Conjuchem, Inc.
;  APPLICANT:  Bridon, Dominique
;  APPLICANT:  Ezrin, Alan
;  APPLICANT:  Milner, Peter
;  APPLICANT:  Holmes, Darren
;  APPLICANT:  Thibaudeau, Karen
;  TITLE OF INVENTION:  PROTECTION OF ENDOGENOUS THERAPEUTIC PEPTIDES FROM
;  TITLE OF INVENTION:  PEPTIDASE ACTIVITY THROUGH CONJUGATION TO BLOOD
;  TITLE OF INVENTION:  COMPONENTS
;  FILE REFERENCE:  2110
;  CURRENT APPLICATION NUMBER: US/09/657,276
;  CURRENT FILING DATE:  2000-09-07
;  PRIOR APPLICATION NUMBER:  60/134,406
;  PRIOR FILING DATE:  1999-05-17
;  PRIOR APPLICATION NUMBER:  60/153,406
;  PRIOR FILING DATE:  1999-09-10
;  PRIOR APPLICATION NUMBER:  60/159,783
;  PRIOR FILING DATE:  1999-10-18
;  NUMBER OF SEQ ID NOS:  1617
;  SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 767
;   LENGTH: 6
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Synthetic
;   OTHER INFORMATION: Peptide
US-09-657-276-767


  Query Match             60.3%;  Score 35;  DB 20;  Length 6;
  Best Local Similarity   83.3%;  Pred. No. 2.6e+06;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       4 PFPGPL 9
           |||||:
Db       1 PFPGPI 6

RESULT   5
US-08-247-451-6
; Sequence 6, Application US/08247451
;  GENERAL INFORMATION:
;    APPLICANT:  ANANTHANARAYANAN, V. S.
;    TITLE OF INVENTION:  PHARMACEUTICAL COMPOSITION AND METHOD
;    TITLE OF INVENTION:  FOR MEDIATING THE PHYSIOLOGICAL EFFECTS OF A COMPOUND
;    NUMBER OF SEQUENCES:  25
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Browdy and Neimark
;      STREET:  419 Seventh St., Suite 300
;      CITY:  Washington, N.W.
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20004
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/247,451
;      FILING DATE:
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/802,982
;      FILING DATE:  06-DEC-1991
;      APPLICATION NUMBER:  US 07/323,421
;      FILING DATE:  14-MAR-1989
;    ATTORNEY/AGENT INFORMATION:
```

```
                         GenCore version 4.5
                Copyright (c) 1993 - 2000  Compugen Ltd.

OM protein - protein search, using sw model

Run on:         October 23, 2001, 13:33:48 ; Search time 81.84 Seconds
                                            (without alignments)
                                            27.107 Million cell updates/sec

Title:          US-09-625-963-3
Perfect score:  58
Sequence:       1 HLMPFPGPLL 10

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched:       2803329 seqs, 221847457 residues

Total number of hits satisfying chosen parameters:      195673

Minimum DB seq length: 0
Maximum DB seq length: 10

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       Pending_Patents_AA_Main:*
               1:  /cgn2_6/ptodata/2/paa/PCTUS_COMB.pep:*
               2:  /cgn2_6/ptodata/2/paa/US06_COMB.pep:*
               3:  /cgn2_6/ptodata/2/paa/US07_COMB.pep:*
               4:  /cgn2_6/ptodata/2/paa/US080_COMB.pep:*
               5:  /cgn2_6/ptodata/2/paa/US081_COMB.pep:*
               6:  /cgn2_6/ptodata/2/paa/US082_COMB.pep:*
               7:  /cgn2_6/ptodata/2/paa/US083_COMB.pep:*
               8:  /cgn2_6/ptodata/2/paa/US084_COMB.pep:*
               9:  /cgn2_6/ptodata/2/paa/US085_COMB.pep:*
               10:  /cgn2_6/ptodata/2/paa/US086_COMB.pep:*
               11:  /cgn2_6/ptodata/2/paa/US087_COMB.pep:*
               12:  /cgn2_6/ptodata/2/paa/US088_COMB.pep:*
               13:  /cgn2_6/ptodata/2/paa/US089_COMB.pep:*
               14:  /cgn2_6/ptodata/2/paa/US090_COMB.pep:*
               15:  /cgn2_6/ptodata/2/paa/US091_COMB.pep:*
               16:  /cgn2_6/ptodata/2/paa/US092_COMB.pep:*
               17:  /cgn2_6/ptodata/2/paa/US093_COMB.pep:*
               18:  /cgn2_6/ptodata/2/paa/US094_COMB.pep:*
               19:  /cgn2_6/ptodata/2/paa/US095_COMB.pep:*
               20:  /cgn2_6/ptodata/2/paa/US096_COMB.pep:*
               21:  /cgn2_6/ptodata/2/paa/US097_COMB.pep:*
               22:  /cgn2_6/ptodata/2/paa/US098_COMB.pep:*
               23:  /cgn2_6/ptodata/2/paa/US60_COMB.pep:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                                SUMMARIES
                   %
Result           Query
  No.   Score  Match Length DB  ID                          Description
-------------------------------------------------------------------------------
    1      58  100.0     10  20  US-09-625-963-3             Sequence 3, Appli
    2      35   60.3      6  17  US-09-317-702-6             Sequence 6, Appli
    3      35   60.3      6  20  US-09-657-276-766           Sequence 766, App
    4      35   60.3      6  20  US-09-657-276-767           Sequence 767, App
    5      35   60.3      7  6  US-08-247-451-6             Sequence 6, Appli
    6      35   60.3      7  17  US-09-317-702-1             Sequence 1, Appli
    7      35   60.3      7  18  US-09-428-692-5             Sequence 5, Appli
    8      35   60.3      7  20  US-09-657-276-748           Sequence 748, App
    9      33   56.9      6  17  US-09-317-702-2             Sequence 2, Appli
   10      33   56.9      9  8  US-08-461-462-8             Sequence 8, Appli
   11      33   56.9      9  8  US-08-461-489-8             Sequence 8, Appli
   12      33   56.9      9  8  US-08-461-788-8             Sequence 8, Appli
   13      33   56.9      9  8  US-08-464-505-8             Sequence 8, Appli
   14      31   53.4      9  4  US-08-027-146-196           Sequence 196, App
   15      31   53.4      9  4  US-08-073-205-196           Sequence 196, App
   16      31   53.4      9  5  US-08-159-184A-196          Sequence 196, App
   17      31   53.4      9  6  US-08-205-713A-196          Sequence 196, App
   18      31   53.4      9  6  US-08-205-713C-196          Sequence 196, App
   19      31   53.4      9  6  US-08-234-784B-27           Sequence 27, Appl
   20      31   53.4      9  10  US-08-682-768-27            Sequence 27, Appl
   21      31   53.4      9  12  US-08-816-454-27            Sequence 27, Appl
   22      31   53.4      9  12  US-08-816-454A-27           Sequence 27, Appl
   23      29   50.0     10  1  PCT-US00-26324-110          Sequence 110, App
   24      28   48.3      7  12  US-08-861-153-24            Sequence 24, Appl
   25      28   48.3      8  11  US-08-781-296-35            Sequence 35, Appl
   26      28   48.3      8  17  US-09-396-347B-33           Sequence 33, Appl
   27      27   46.6      7  12  US-08-861-153-22            Sequence 22, Appl
   28      27   46.6      7  12  US-08-861-153-23            Sequence 23, Appl
   29      27   46.6      9  6  US-08-234-784B-26           Sequence 26, Appl
   30      27   46.6      9  10  US-08-682-768-26            Sequence 26, Appl
   31      27   46.6      9  12  US-08-816-454-26            Sequence 26, Appl
   32      27   46.6      9  12  US-08-816-454A-26           Sequence 26, Appl
   33      27   46.6      9  14  US-09-073-138-111           Sequence 111, App
   34      27   46.6      9  14  US-09-073-138-119           Sequence 119, App
   35      27   46.6      9  16  US-09-267-439-111           Sequence 111, App
   36      27   46.6      9  16  US-09-267-439-119           Sequence 119, App
   37      27   46.6      9  19  US-09-529-206-65            Sequence 65, Appl
   38      27   46.6      9  20  US-09-668-143-14            Sequence 14, Appl
   39      27   46.6     10  1  PCT-US00-22665-18           Sequence 18, Appl
   40      27   46.6     10  1  PCT-US00-22774-18           Sequence 18, Appl
   41      27   46.6     10  1  PCT-US00-22818-18           Sequence 18, Appl
   42      27   46.6     10  6  US-08-234-784B-69           Sequence 69, Appl
   43      27   46.6     10  10  US-08-682-768-69            Sequence 69, Appl
   44      27   46.6     10  12  US-08-816-454-69            Sequence 69, Appl
   45      27   46.6     10  12  US-08-816-454A-69           Sequence 69, Appl


                                ALIGNMENTS


RESULT   1
US-09-625-963-3
; Sequence 3, Application US/09625963
; GENERAL INFORMATION:
;  APPLICANT:  Strauss, Hans Josef
;  APPLICANT:  Gao, Liquan
;  TITLE OF INVENTION:  Immunotherapeutic Methods Using Epitopes of WT-1 and
;  TITLE OF INVENTION:  GATA-1
;  FILE REFERENCE:  ICI 101
;  CURRENT APPLICATION NUMBER: US/09/625,963
;  CURRENT FILING DATE:  2000-07-26
;  PRIOR APPLICATION NUMBER:  PCT/GB99/03572
;  PRIOR FILING DATE:  1999-11-02
;  PRIOR APPLICATION NUMBER:  GB9823897.5
;  PRIOR FILING DATE:  1998-11-02
;  NUMBER OF SEQ ID NOS:  5
;  SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 3
;   LENGTH: 10
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-625-963-3


  Query Match             100.0%;  Score 58;  DB 20;  Length 10;
  Best Local Similarity   100.0%;  Pred. No. 0.038;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 HLMPFPGPLL 10
              ||||||||||
Db          1 HLMPFPGPLL 10
```

```
PR   07-NOV-1991;   91JP-0318569.
PR   25-DEC-1991;   91JP-0356633.
PR   01-SEP-1992;   92JP-0255403.
XX
PA   (NISS ) NISSHIN FLOUR MILLING CO.
XX
DR   WPI; 1994-188987/23.
XX
PT   New opioid peptide(s) and their salts - are used as analgesic,
PT   for stimulating e.g. electrolyte absorption and treating
PT   diarrhoea
XX
PS   Disclosure; Page 13; 15pp; Japanese.
XX
CC   The sequences given in AAR53607-26 are peptides which have opioid
CC   activity.  These peptides are produced by hydrolysing lactoprotein
CC   with neutral protease derived from alkali protease.  The peptides and
CC   their salts are useful as analgesics, hypnotics, gastrointestinal
CC   hormone stimulating agents, electrolyte adsorption stimulating agents,
CC   diarrhoea treating agents, etc.  The peptide given in AAR53608 shows
CC   lower opioid activity and is therefore expected to be used as processed
CC   drug.
XX
SQ   Sequence   8 AA;


  Query Match             60.3%;  Score 35;  DB 15;  Length 8;
  Best Local Similarity   83.3%;  Pred. No. 3.4e+05;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        4 PFPGPL 9
            |||||:
Db        3 pfpgpi 8


RESULT 15
AAR26793
ID   AAR26793 standard; peptide; 9 AA.
XX
AC   AAR26793;
XX
DT   11-FEB-1993  (first entry)
XX
DE   Prolyl endopeptidase inhibitor (1).
XX
KW   Amnesia; endopeptidase.
XX
OS   Synthetic.
XX
PN   JP04208299-A.
XX
PD   29-JUL-1992.
XX
PF   30-NOV-1990;   90JP-0337219.
XX
PR   30-NOV-1990;   90JP-0337219.
XX
PA   (AJIN ) AJINOMOTO KK.
XX
DR   WPI; 1992-303600/37.
XX
PT   Prolyl endo:peptidase inhibitors - for treating and preventing
PT   amnesia
XX
PS   Claim 1; Page 1; 11pp; Japanese.
XX
CC   The peptides given in AAR26793-801 inhibit prolyl endopeptidase and
CC   are used in prevention and treatment of amnesia.  They may be
CC   administered orally as powder, granules, tablets, capsules or liq.
CC   prepn. or parenterally as suspension, liq. prepn., emulsion or
CC   injection.  E.g. they may be included as a component of a nutrient.
CC   The peptides may be prepd. by condensation of two suitable fragments,
CC   i.e. one having a reactive carboxy gp. and the other having
CC   a reactive amino gp., by a method using DCC, and subsequent
CC   deprotection.  The condensation may be made by the liq.-phase or
CC   solid-phase method.
XX
SQ   Sequence   9 AA;

  Query Match             60.3%;  Score 35;  DB 13;  Length 9;
  Best Local Similarity   83.3%;  Pred. No. 3.4e+05;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        4 PFPGPL 9
            |||||:
Db        3 pfpgpi 8


Search completed: October 23, 2001, 13:34:29
Job time: 276 sec
```

```
Qy       4 PFPGPL 9
           |||||:
Db       2 pfpgpi 7


RESULT  12
AAY49371
ID   AAY49371 standard; peptide; 7 AA.
XX
AC   AAY49371;
XX
DT   13-MAR-2000  (first entry)
XX
DE   Beta-casomorphin derivative peptide as diagnostic marker.
XX
KW   Diagnostic marker; human disorder; opiate; autism spectral disorder;
KW   autism pervasive developmental disorder; Aspergers syndrome;
KW   attention deficient disorder; attention hyperactivity disorder;
KW   multiple sclerosis; Parkinson's disease; Alzheimer's dementia;
KW   beta-casomorphin.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
PN   EP969015-A2.
XX
PD   05-JAN-2000.
XX
PF   15-JUN-1999;   99EP-0304636.
XX
PR   15-JUN-1998;   98US-0089237.
PR   15-JUN-1998;   98US-0089238.
PR   24-MAY-1999;   99US-0317702.
XX
PA   (ORTH-) ORTHO-CLINICAL DIAGNOSTICS INC.
XX
PI   Shanahan MR,  Venturini AJ,  Daiss JL,  Friedman AE;
XX
DR   WPI; 2000-074781/07.
XX
PT   Diagnosing human disorders e.g. autism spectral disorders, multiple
PT   sclerosis, Parkinson's disease and Alzheimer's dementia -
XX
PS   Claim 4; Page 6; 44pp; English.
XX
CC   The invention provides diagnostic markers for a human disorder,
CC   comprising either opiate-like peptides or opiate-derived peptides. The
CC   novel peptides are used as diagnostic marker, in ex-vivo methods of
CC   diagnosing human disorders e.g. autism spectral disorders including
CC   autism pervasive developmental disorder, Aspergers syndrome, attention
CC   deficient disorder and attention hyperactivity disorder, and multiple
CC   sclerosis, Parkinson's disease and Alzheimer's dementia. Sequences
CC   AAY49371-413 represent peptide diagnostic markers for the human
CC   disorders specified above.
XX
SQ   Sequence   7 AA;

  Query Match             60.3%;  Score 35;  DB 21;  Length 7;
  Best Local Similarity   83.3%;  Pred. No. 3.4e+05;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       4 PFPGPL 9
           |||||:
Db       2 pfpgpi 7


RESULT  13
AAB74637
ID   AAB74637 standard; peptide; 7 AA.
XX
AC   AAB74637;
XX
DT   31-MAY-2001  (first entry)
XX
DE   Milk protein fragment.
XX
KW   Milk protein; mammalian; central nervous system; tranquilliser;
KW   behavioural; neurochemical; anxiety; anxiolytic.
XX
OS   Mammalia.
XX
PN   RU2161500-C1.
XX
PD   10-JAN-2001.
XX
PF   06-JUL-1999;   99RU-0114696.
XX
PR   06-JUL-1999;   99RU-0114696.
XX
PA   (ASMO=) AS USSR MOLECULAR GENETICS INST.
XX
PI   Dubynin VA,  Ashmarin IP,  Kamenskii AA,  Malinovskaya IV;
PI   Alfeeva Yu L,  Andreeva LA,  Myasoedov NF;
XX
DR   WPI; 2001-217067/22.
XX
PT   Heptapeptide from milk protein used for reducing anxiety and increasing
PT   central nervous system capacity  -
XX
PS   Claim 1; Column 5; 5pp; Russian.
XX
CC   The present invention describes a method which involves applying a milk
CC   protein fragment (the present heptapeptide sequence:
CC   Tyr-Pro-Phe-Pro-Gly-pro-Ile) having anxiolytic activity to mammalian
CC   youth proved in behavioural and neurochemical tests carried out 1-7
CC   weeks later after cancelling peptide administration. The heptapeptide
CC   has tranquilliser and central nervous system activities. The method is
CC   useful for reducing anxiety and increasing central nervous system
CC   capability in mammals.
XX
SQ   Sequence   7 AA;

  Query Match             60.3%;  Score 35;  DB 22;  Length 7;
  Best Local Similarity   83.3%;  Pred. No. 3.4e+05;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       4 PFPGPL 9
           |||||:
Db       2 pfpgpi 7


RESULT  14
AAR53626
ID   AAR53626 standard; peptide; 8 AA.
XX
AC   AAR53626;
XX
DT   25-JAN-1995  (first entry)
XX
DE   Opioid peptide #20.
XX
KW   Opioid; activity; hydrolysis; lactoprotein; neutral protease;
KW   alkali protease; gastrointestinal hormone stimulating agents;
KW   analgesics; hypnotics; electrolyte adsorption stimulating agents;
KW   diarrhoea treating agents; drug.
XX
OS   Mammalian.
XX
PN   JP06128287-A.
XX
PD   10-MAY-1994.
XX
PF   02-NOV-1992;   92JP-0315515.
XX
```

```
XX
PN   WO9215279-A.
XX
PD   17-SEP-1992.
XX
PF   02-MAR-1992;   92WO-FR00191.
XX
PR   01-MAR-1991;   91FR-0002510.
XX
PA   (OREA ) L'OREAL SA.
XX
PI   Barey M,  Bitri L,  Darmanton P;
XX
DR   WPI; 1992-331438/40.
XX
PT   Cosmetic compsn. contg. peptide as active principle - obtd. by
PT   hydrolysis of casein, with opioid activity, used for improving
PT   appearance of skin sunburn or other irritation
XX
PS   Claim 1; Page 13; 16pp; French.
XX
CC   The inventors claim a cosmetic compsn. contg. the peptide as an
CC   active ingredient. The peptide is derived from milk casein and has
CC   opioid activity. Compsns. contg. it are used to improve the
CC   appearance of sunburned skin or skin irritated by other agents or to
CC   counteract the effects of ageing on the skin. The amt. of peptide in
CC   the compsn. is 0.1-10 (pref. 0.1-5)a wt. %. The peptide is free of
CC   other peptides of mol. wt. > 3000.
XX
SQ   Sequence   7 AA;

  Query Match             60.3%;  Score 35;  DB 13;  Length 7;
  Best Local Similarity   83.3%;  Pred. No. 3.4e+05;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       4 PFPGPL 9
           |||||:
Db       2 pfpgpi 7

RESULT  10
AAR53416
ID   AAR53416 standard; peptide; 7 AA.
XX
AC   AAR53416;
XX
DT   30-DEC-1994  (first entry)
XX
DE   Peptide having opioid activity.
XX
KW   Opioid; analgesic; sleeping drug;
KW   gastrointestinal hormone secretion promoter;
KW   electrolyte absorption promoter; diarrhoea.
XX
OS   Synthetic.
XX
PN   JP06107684-A.
XX
PD   19-APR-1994.
XX
PF   02-NOV-1992;   92JP-0315516.
XX
PR   14-AUG-1992;   92JP-0237577.
XX
PA   (NISS ) NISSHIN FLOUR MILLING CO.
XX
DR   WPI; 1994-163937/20.
XX
PT   New penta:peptide(s) having opioid activity - useful e.g. as
PT   analgesics, sleeping drugs, gastrointestinal hormone secretion
PT   promoters and electrolyte absorption promoters
XX
```

```
PS   Claim 1; Page 5; 6pp; Japanese.
XX
CC   The peptide has high opioid activity, being useful as an analgesic,
CC   sleeping drug, gastrointestinal hormone secretion promoter,
CC   electrolyte absorption promoter and as an agent for mitigating
CC   diarrhoea due to constriction of the gastrointestinal transport
CC   muscle.
XX
SQ   Sequence   7 AA;

  Query Match             60.3%;  Score 35;  DB 15;  Length 7;
  Best Local Similarity   83.3%;  Pred. No. 3.4e+05;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       4 PFPGPL 9
           |||||:
Db       2 pfpgpi 7

RESULT  11
AAR53623
ID   AAR53623 standard; peptide; 7 AA.
XX
AC   AAR53623;
XX
DT   25-JAN-1995  (first entry)
XX
DE   Opioid peptide #17.
XX
KW   Opioid; activity; hydrolysis; lactoprotein; neutral protease;
KW   alkali protease; gastrointestinal hormone stimulating agents;
KW   analgesics; hypnotics; electrolyte adsorption stimulating agents;
KW   diarrhoea treating agents; drug.
XX
OS   Mammalian.
XX
PN   JP06128287-A.
XX
PD   10-MAY-1994.
XX
PF   02-NOV-1992;   92JP-0315515.
XX
PR   07-NOV-1991;   91JP-0318569.
PR   25-DEC-1991;   91JP-0356633.
PR   01-SEP-1992;   92JP-0255403.
XX
PA   (NISS ) NISSHIN FLOUR MILLING CO.
XX
DR   WPI; 1994-188987/23.
XX
PT   New opioid peptide(s) and their salts - are used as analgesic,
PT   for stimulating e.g. electrolyte absorption and treating
PT   diarrhoea
XX
PS   Claim 10; Page 13; 15pp; Japanese.
XX
CC   The sequences given in AAR53607-26 are peptides which have opioid
CC   activity.  These peptides are produced by hydrolysing lactoprotein
CC   with neutral protease derived from alkali protease.  The peptides and
CC   their salts are useful as analgesics, hypnotics, gastrointestinal
CC   hormone stimulating agents, electrolyte adsorption stimulating agents,
CC   diarrhoea treating agents, etc.  The peptide given in AAR53608 shows
CC   lower opioid activity and is therefore expected to be used as processed
CC   drug.
XX
SQ   Sequence   7 AA;

  Query Match             60.3%;  Score 35;  DB 15;  Length 7;
  Best Local Similarity   83.3%;  Pred. No. 3.4e+05;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
  Matches     5;  Conservative     2;  Mismatches     1;  Indels     0;  Gaps     0;

Qy       2 LMPFPGPL 9
           : |||||:
Db       1 iypfpgpi 8


RESULT   7
AAY49376
ID   AAY49376 standard; peptide; 6 AA.
XX
AC   AAY49376;
XX
DT   13-MAR-2000  (first entry)
XX
DE   Beta-casomorphin derivative peptide as diagnostic marker.
XX
KW   Diagnostic marker; human disorder; opiate; autism spectral disorder;
KW   autism pervasive developmental disorder; Aspergers syndrome;
KW   attention deficient disorder; attention hyperactivity disorder;
KW   multiple sclerosis; Parkinson's disease; Alzheimer's dementia;
KW   beta-casomorphin.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
PN   EP969015-A2.
XX
PD   05-JAN-2000.
XX
PF   15-JUN-1999;   99EP-0304636.
XX
PR   15-JUN-1998;   98US-0089237.
PR   15-JUN-1998;   98US-0089238.
PR   24-MAY-1999;   99US-0317702.
XX
PA   (ORTH-) ORTHO-CLINICAL DIAGNOSTICS INC.
XX
PI   Shanahan MR,  Venturini AJ,  Daiss JL,  Friedman AE;
XX
DR   WPI; 2000-074781/07.
XX
PT   Diagnosing human disorders e.g. autism spectral disorders, multiple
PT   sclerosis, Parkinson's disease and Alzheimer's dementia -
XX
PS   Claim 4; Page 6; 44pp; English.
XX
CC   The invention provides diagnostic markers for a human disorder,
CC   comprising either opiate-like peptides or opiate-derived peptides. The
CC   novel peptides are used as diagnostic marker, in ex-vivo methods of
CC   diagnosing human disorders e.g. autism spectral disorders including
CC   autism pervasive developmental disorder, Aspergers syndrome, attention
CC   deficient disorder and attention hyperactivity disorder, and multiple
CC   sclerosis, Parkinson's disease and Alzheimer's dementia. Sequences
CC   AAY49371-413 represent peptide diagnostic markers for the human
CC   disorders specified above.
XX
SQ   Sequence   6 AA;


  Query Match             60.3%;  Score 35;  DB 21;  Length 6;
  Best Local Similarity   83.3%;  Pred. No. 3.4e+05;
  Matches     5;  Conservative     1;  Mismatches     0;  Indels     0;  Gaps     0;

Qy       4 PFPGPL 9
             |||||:
Db       1 pfpgpi 6


RESULT   8
AAP50912
ID   AAP50912 standard; peptide; 7 AA.
XX
AC   AAP50912;
XX
DT   30-NOV-1991  (first entry)
XX
DE   Sequence of morphiceptin and beta-casomorphin used in
DE   radioimmunoassay for determining the risk of sudden infant death
DE   syndrome.
XX
KW   Cot death; radioimmunoassay; sudden infant death syndrome.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /label= H-(125I)iodo-Tyr
FT   Modified-site   7
FT                   /label= Ile-OH
FT   Misc-difference 5..7
FT                   /note= "absent from morphiceptin"
FT   Modified-site   4
FT                   /label= Pro-NH2
FT                   /note= "morphiceptin only"
XX
PN   EP139201-A.
XX
PD   02-MAY-1985.
XX
PF   07-SEP-1984;   84EP-0110703.
XX
PR   09-SEP-1983;   83GB-0024167.
XX
PA   (WELL ) WELLCOME FOUNDATION LTD.
XX
PI   Chang KJ;
XX
DR   WPI; 1985-106364/18.
XX
PT   Determining the risk of sudden infant death syndrome - by
PT   measuring concn. of morphiceptin or beta-casomorphin in the blood
XX
PS   Claim 16; Page 13; 15pp; English.
XX
CC   AAs 1-4 represent morphiceptin; AAs 1-7 represent beta-casomorphin.
CC   The assay test sample, pref. plasma or serum, is incubated with
CC   labelled peptides and specific antibody (immobilised on protein A-
CC   contg. Staphylococcus), pref. in a buffer. The liq. and solid phases
CC   are then sepd. and the gamma activity of the latter counted.
XX
SQ   Sequence   7 AA;


  Query Match             60.3%;  Score 35;  DB 6;  Length 7;
  Best Local Similarity   83.3%;  Pred. No. 3.4e+05;
  Matches     5;  Conservative     1;  Mismatches     0;  Indels     0;  Gaps     0;

Qy       4 PFPGPL 9
             |||||:
Db       2 pfpgpi 7


RESULT   9
AAR27171
ID   AAR27171 standard; Protein; 7 AA.
XX
AC   AAR27171;
XX
DT   03-FEB-1993  (first entry)
XX
DE   Sequence of peptide derived from milk casein with opioid activity.
XX
KW   Milk casein; opioid; sunburn; ageing; therapy.
XX
OS   Synthetic.
```

```
XX
KW   Amnesia; endopeptidase.
XX
OS   Synthetic.
XX
PN   JP04208299-A.
XX
PD   29-JUL-1992.
XX
PF   30-NOV-1990;   90JP-0337219.
XX
PR   30-NOV-1990;   90JP-0337219.
XX
PA   (AJIN ) AJINOMOTO KK.
XX
DR   WPI; 1992-303600/37.
XX
PT   Prolyl endo:peptidase inhibitors - for treating and preventing
PT   amnesia
XX
PS   Claim 1; Page 1; 11pp; Japanese.
XX
CC   The peptides given in AAR26793-801 inhibit prolyl endopeptidase and
CC   are used in prevention and treatment of amnesia.  They may be
CC   administered orally as powder, granules, tablets, capsules or liq.
CC   prepn. or parenterally as suspension, liq. prepn., emulsion or
CC   injection.  E.g. they may be included as a component of a nutrient.
CC   The peptides may be prepd. by condensation of two suitable fragments,
CC   i.e. one having a reactive carboxy gp. and the other having
CC   a reactive amino gp., by a method using DCC, and subsequent
CC   deprotection.  The condensation may be made by the liq.-phase or
CC   solid-phase method.
XX
SQ   Sequence   8 AA;


  Query Match             62.1%;  Score 36;  DB 13;  Length 8;
  Best Local Similarity   62.5%;  Pred. No. 3.4e+05;
  Matches    5;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      2 LMPFPGPL 9
          : |||||:
Db      1 iypfpgpi 8


RESULT   5
AAR26798
ID   AAR26798 standard; peptide; 9 AA.
XX
AC   AAR26798;
XX
DT   11-FEB-1993  (first entry)
XX
DE   Prolyl endopeptidase inhibitor (6).
XX
KW   Amnesia; endopeptidase.
XX
OS   Synthetic.
XX
PN   JP04208299-A.
XX
PD   29-JUL-1992.
XX
PF   30-NOV-1990;   90JP-0337219.
XX
PR   30-NOV-1990;   90JP-0337219.
XX
PA   (AJIN ) AJINOMOTO KK.
XX
DR   WPI; 1992-303600/37.
XX
PT   Prolyl endo:peptidase inhibitors - for treating and preventing
PT   amnesia
```

```
XX
PS   Claim 1; Page 1; 11pp; Japanese.
XX
CC   The peptides given in AAR26793-801 inhibit prolyl endopeptidase and
CC   are used in prevention and treatment of amnesia.  They may be
CC   administered orally as powder, granules, tablets, capsules or liq.
CC   prepn. or parenterally as suspension, liq. prepn., emulsion or
CC   injection.  E.g. they may be included as a component of a nutrient.
CC   The peptides may be prepd. by condensation of two suitable fragments,
CC   i.e. one having a reactive carboxy gp. and the other having
CC   a reactive amino gp., by a method using DCC, and subsequent
CC   deprotection.  The condensation may be made by the liq.-phase or
CC   solid-phase method.
XX
SQ   Sequence   9 AA;


  Query Match             62.1%;  Score 36;  DB 13;  Length 9;
  Best Local Similarity   62.5%;  Pred. No. 3.4e+05;
  Matches    5;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      2 LMPFPGPL 9
          : |||||:
Db      1 iypfpgpi 8


RESULT   6
AAR26797
ID   AAR26797 standard; peptide; 9 AA.
XX
AC   AAR26797;
XX
DT   11-FEB-1993  (first entry)
XX
DE   Prolyl endopeptidase inhibitor (5).
XX
KW   Amnesia; endopeptidase.
XX
OS   Synthetic.
XX
PN   JP04208299-A.
XX
PD   29-JUL-1992.
XX
PF   30-NOV-1990;   90JP-0337219.
XX
PR   30-NOV-1990;   90JP-0337219.
XX
PA   (AJIN ) AJINOMOTO KK.
XX
DR   WPI; 1992-303600/37.
XX
PT   Prolyl endo:peptidase inhibitors - for treating and preventing
PT   amnesia
XX
PS   Claim 1; Page 1; 11pp; Japanese.
XX
CC   The peptides given in AAR26793-801 inhibit prolyl endopeptidase and
CC   are used in prevention and treatment of amnesia.  They may be
CC   administered orally as powder, granules, tablets, capsules or liq.
CC   prepn. or parenterally as suspension, liq. prepn., emulsion or
CC   injection.  E.g. they may be included as a component of a nutrient.
CC   The peptides may be prepd. by condensation of two suitable fragments,
CC   i.e. one having a reactive carboxy gp. and the other having
CC   a reactive amino gp., by a method using DCC, and subsequent
CC   deprotection.  The condensation may be made by the liq.-phase or
CC   solid-phase method.
XX
SQ   Sequence   9 AA;


  Query Match             62.1%;  Score 36;  DB 13;  Length 9;
  Best Local Similarity   62.5%;  Pred. No. 3.4e+05;
```

```
CC   The present sequence is peptide epitope hug 378-87, produced by gata-1
CC   expressing cells and found at residues 378-387 of the gata-1 protein,
CC   which is recognised by cytotoxic T lymphocytes. Gata-1 is aberrantly
CC   expressed in leukaemias. The peptide can be used as a vaccine to
CC   stimulate the elimination, by cytotoxic T lymphocytes, of cancer cells
CC   aberrantly expressing gata-1. In addition, the nucleic acid encoding the
CC   peptide may also be used in the same manner. Alternatively, the peptide
CC   may be used in vitro to produce activated cytotoxic T lymphocytes.
XX
SQ   Sequence   10 AA;

  Query Match             100.0%;  Score 58;  DB 21;  Length 10;
  Best Local Similarity   100.0%;  Pred. No. 0.00095;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 HLMPFPGPLL 10
            ||||||||||
Db        1 hlmpfpgpll 10


RESULT   2
AAR53617
ID   AAR53617 standard; peptide; 9 AA.
XX
AC   AAR53617;
XX
DT   25-JAN-1995  (first entry)
XX
DE   Opioid peptide #11.
XX
KW   Opioid; activity; hydrolysis; lactoprotein; neutral protease;
KW   alkali protease; gastrointestinal hormone stimulating agents;
KW   analgesics; hypnotics; electrolyte adsorption stimulating agents;
KW   diarrhoea treating agents; drug.
XX
OS   Mammalian.
XX
PN   JP06128287-A.
XX
PD   10-MAY-1994.
XX
PF   02-NOV-1992;   92JP-0315515.
XX
PR   07-NOV-1991;   91JP-0318569.
PR   25-DEC-1991;   91JP-0356633.
PR   01-SEP-1992;   92JP-0255403.
XX
PA   (NISS ) NISSHIN FLOUR MILLING CO.
XX
DR   WPI; 1994-188987/23.
XX
PT   New opioid peptide(s) and their salts - are used as analgesic,
PT   for stimulating e.g. electrolyte absorption and treating
PT   diarrhoea
XX
PS   Disclosure; Page 13; 15pp; Japanese.
XX
CC   The sequences given in AAR53607-26 are peptides which have opioid
CC   activity.  These peptides are produced by hydrolysing lactoprotein
CC   with neutral protease derived from alkali protease.  The peptides and
CC   their salts are useful as analgesics, hypnotics, gastrointestinal
CC   hormone stimulating agents, electrolyte adsorption stimulating agents,
CC   diarrhoea treating agents, etc.  The peptide given in AAR53608 shows
CC   lower opioid activity and is therefore expected to be used as processed
CC   drug.
XX
SQ   Sequence   9 AA;

  Query Match             67.2%;  Score 39;  DB 15;  Length 9;
  Best Local Similarity   85.7%;  Pred. No. 3.4e+05;
  Matches    6;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        4 PFPGPLL 10
              |||||:|
Db        2 pfpgpil 8


RESULT   3
AAR53622
ID   AAR53622 standard; peptide; 10 AA.
XX
AC   AAR53622;
XX
DT   25-JAN-1995  (first entry).
XX
DE   Opioid peptide #16.
XX
KW   Opioid; activity; hydrolysis; lactoprotein; neutral protease;
KW   alkali protease; gastrointestinal hormone stimulating agents;
KW   analgesics; hypnotics; electrolyte adsorption stimulating agents;
KW   diarrhoea treating agents; drug.
XX
OS   Mammalian.
XX
PN   JP06128287-A.
XX
PD   10-MAY-1994.
XX
PF   02-NOV-1992;   92JP-0315515.
XX
PR   07-NOV-1991;   91JP-0318569.
PR   25-DEC-1991;   91JP-0356633.
PR   01-SEP-1992;   92JP-0255403.
XX
PA   (NISS ) NISSHIN FLOUR MILLING CO.
XX
DR   WPI; 1994-188987/23.
XX
PT   New opioid peptide(s) and their salts - are used as analgesic,
PT   for stimulating e.g. electrolyte absorption and treating
PT   diarrhoea
XX
PS   Disclosure; Page 13; 15pp; Japanese.
XX
CC   The sequences given in AAR53607-26 are peptides which have opioid
CC   activity.  These peptides are produced by hydrolysing lactoprotein
CC   with neutral protease derived from alkali protease.  The peptides and
CC   their salts are useful as analgesics, hypnotics, gastrointestinal
CC   hormone stimulating agents, electrolyte adsorption stimulating agents,
CC   diarrhoea treating agents, etc.  The peptide given in AAR53608 shows
CC   lower opioid activity and is therefore expected to be used as processed
CC   drug.
XX
SQ   Sequence   10 AA;

  Query Match             67.2%;  Score 39;  DB 15;  Length 10;
  Best Local Similarity   85.7%;  Pred. No. 1.4;
  Matches    6;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        4 PFPGPLL 10
              |||||:|
Db        3 pfpgpil 9


RESULT   4
AAR26796
ID   AAR26796 standard; peptide; 8 AA.
XX
AC   AAR26796;
XX
DT   11-FEB-1993  (first entry)
XX
DE   Prolyl endopeptidase inhibitor (4).
```

```
                        GenCore version 4.5
              Copyright (c) 1993 - 2000  Compugen Ltd.


OM protein - protein search, using sw model

Run on:         October 23, 2001, 13:29:53 ; Search time 17.77 Seconds
                                             (without alignments)
                                             34.116 Million cell updates/sec

Title:          US-09-625-963-3
Perfect score:  58
Sequence:       1 HLMPFPGPLL 10

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched:       412676 seqs, 60623988 residues

Total number of hits satisfying chosen parameters:      90768

Minimum DB seq length: 0
Maximum DB seq length: 10

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :      A_Geneseq_0601:*
               1:  /SIDS8/gcgdata/geneseq/geneseqp/AA1980.DAT:*
               2:  /SIDS8/gcgdata/geneseq/geneseqp/AA1981.DAT:*
               3:  /SIDS8/gcgdata/geneseq/geneseqp/AA1982.DAT:*
               4:  /SIDS8/gcgdata/geneseq/geneseqp/AA1983.DAT:*
               5:  /SIDS8/gcgdata/geneseq/geneseqp/AA1984.DAT:*
               6:  /SIDS8/gcgdata/geneseq/geneseqp/AA1985.DAT:*
               7:  /SIDS8/gcgdata/geneseq/geneseqp/AA1986.DAT:*
               8:  /SIDS8/gcgdata/geneseq/geneseqp/AA1987.DAT:*
               9:  /SIDS8/gcgdata/geneseq/geneseqp/AA1988.DAT:*
              10:  /SIDS8/gcgdata/geneseq/geneseqp/AA1989.DAT:*
              11:  /SIDS8/gcgdata/geneseq/geneseqp/AA1990.DAT:*
              12:  /SIDS8/gcgdata/geneseq/geneseqp/AA1991.DAT:*
              13:  /SIDS8/gcgdata/geneseq/geneseqp/AA1992.DAT:*
              14:  /SIDS8/gcgdata/geneseq/geneseqp/AA1993.DAT:*
              15:  /SIDS8/gcgdata/geneseq/geneseqp/AA1994.DAT:*
              16:  /SIDS8/gcgdata/geneseq/geneseqp/AA1995.DAT:*
              17:  /SIDS8/gcgdata/geneseq/geneseqp/AA1996.DAT:*
              18:  /SIDS8/gcgdata/geneseq/geneseqp/AA1997.DAT:*
              19:  /SIDS8/gcgdata/geneseq/geneseqp/AA1998.DAT:*
              20:  /SIDS8/gcgdata/geneseq/geneseqp/AA1999.DAT:*
              21:  /SIDS8/gcgdata/geneseq/geneseqp/AA2000.DAT:*
              22:  /SIDS8/gcgdata/geneseq/geneseqp/AA2001.DAT:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                                  SUMMARIES
                 %
Result           Query
  No.    Score   Match Length DB  ID                       Description
  -----------------------------------------------------------------------
     1      58   100.0     10  21  AAY94204                 Human cytotoxic T
     2      39    67.2      9  15  AAR53617                 Opioid peptide #11
     3      39    67.2     10  15  AAR53622                 Opioid peptide #16
     4      36    62.1      8  13  AAR26796                 Prolyl endopeptida
     5      36    62.1      9  13  AAR26798                 Prolyl endopeptida
     6      36    62.1      9  13  AAR26797                 Prolyl endopeptida
     7      35    60.3      6  21  AAY49376                 Beta-casomorphin d
     8      35    60.3      7  6  AAP50912                 Sequence of morphi
     9      35    60.3      7  13  AAR27171                 Sequence of peptid
    10      35    60.3      7  15  AAR53416                 Peptide having opi
    11      35    60.3      7  15  AAR53623                 Opioid peptide #17
```

```
    12      35    60.3      7  21  AAY49371                 Beta-casomorphin d
    13      35    60.3      7  22  AAB74637                 Milk protein fragm
    14      35    60.3      8  15  AAR53626                 Opioid peptide #20
    15      35    60.3      9  13  AAR26793                 Prolyl endopeptida
    16      35    60.3      9  13  AAR26794                 Prolyl endopeptida
    17      35    60.3      9  13  AAR26795                 Prolyl endopeptida
    18      35    60.3      9  15  AAR53613                 Opioid peptide #7.
    19      35    60.3      9  15  AAR53614                 Opioid peptide #8.
    20      35    60.3      9  15  AAR53615                 Opioid peptide #9.
    21      35    60.3      9  15  AAR53616                 Opioid peptide #10
    22      35    60.3      9  15  AAR53611                 Opioid peptide #5.
    23      35    60.3      9  15  AAR53612                 Opioid peptide #6.
    24      35    60.3     10  15  AAR53618                 Opioid peptide #12
    25      35    60.3     10  15  AAR53619                 Opioid peptide #13
    26      35    60.3     10  15  AAR53620                 Opioid peptide #14
    27      35    60.3     10  15  AAR53621                 Opioid peptide #15
    28      35    60.3     10  15  AAR53608                 Opioid peptide #2.
    29      35    60.3     10  15  AAR53609                 Opioid peptide #3.
    30      35    60.3     10  15  AAR53610                 Opioid peptide #4.
    31      33    56.9      6  21  AAY49372                 Beta-casomorphin d
    32      31    53.4      9  15  AAR59241                 Peptide fragment (
    33      31    53.4      9  16  AAR82133                 Melanoma-specific
    34      30    51.7      9  8  AAP70197                 Sequence of bovine
    35      30    51.7      9  16  AAR76504                 Cattle MIS tryptic
    36      29    50.0      7  19  AAW47467                 Prothrombin/thromb
    37      29    50.0      9  17  AAW49253                 Human leucocyte an
    38      28    48.3      7  20  AAW74233                 HJ loop peptide HJ
    39      28    48.3      8  19  AAW70136                 Peptide produced b
    40      28    48.3      8  19  AAW65584                 Epstein-Barr virus
    41      28    48.3      9  20  AAY10112                 T cell epitope/MHC
    42      27    46.6      6  21  AAY67446                 SH3 domain binding
    43      27    46.6      7  20  AAW74229                 HJ loop peptide HJ
    44      27    46.6      7  20  AAW74231                 HJ loop peptide HJ
    45      27    46.6      9  16  AAR82132                 Melanoma-specific


                              ALIGNMENTS


RESULT   1
AAY94204
ID   AAY94204 standard; peptide; 10 AA.
XX
AC   AAY94204;
XX
DT   28-JUL-2000  (first entry)
XX
DE   Human cytotoxic T lymphocyte-recognised gata-1 peptide hug 378-87.
XX
KW   Hug 378-87; peptide; epitope; gata-1 gene; leukaemia; immunotherapy;
KW   transcription factor.
XX
OS   Homo sapiens.
XX
PN   WO200026249-A1.
XX
PD   11-MAY-2000.
XX
PF   02-NOV-1999;   99WO-GB03572.
XX
PR   02-NOV-1998;   98GB-0023897.
XX
PA   (IMCO-) IMPERIAL COLLEGE INNOVATIONS LTD.
XX
PI   Stauss HJ,  Gao L;
XX
DR   WPI; 2000-376123/32.
XX
PT   Novel peptides comprising WT-1 and GATA-1 epitopes, their fragments or
PT   variants, useful as vaccines for cancer immunotherapy  -
XX
PS   Claim 3; Page 74; 93pp; English.
XX
```

```
ID   O93739      PRELIMINARY;   PRT;   10 AA.
AC   O93739;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAY-1999 (TrEMBLrel. 10, Last annotation update)
DE   7S RNA PROTEIN (FRAGMENT).
GN   7S RNA.
OS   Methanothermus fervidus.
OC   Archaea; Euryarchaeota; Methanobacteriales; Methanothermaceae;
OC   Methanothermus.
OX   NCBI_TaxID=2180;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93093426; PubMed=1281131;
RA   Koller G., Reeve J.N., Frey G., Thomm M.;
RT   "Transcription in vitro and in vivo of the 7S RNA gene associated with
RT   the ribosomal RNA operon in the hyperthermophilic archaeon
RT   Methanothermus fervidus.";
RL   FEMS Microbiol. Lett. 77:95-101(1992).
DR   EMBL; S49762; AAD13859.1; -.
FT   NON_TER       1      1
SQ   SEQUENCE   10 AA;  1080 MW;  220CF26412C86408 CRC64;


  Query Match             29.3%;  Score 17;  DB 1;  Length 10;
  Best Local Similarity   100.0%;  Pred. No. 8.2e+03;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      7 GPL 9
          |||
Db      8 GPL 10


Search completed: October 23, 2001, 13:37:26
Job time: 173 sec
```

```
  Query Match             31.0%;  Score 18;  DB 11;  Length 10;
  Best Local Similarity   60.0%;  Pred. No. 5.6e+03;
  Matches    3;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       4 PFPGP 8
           |  ||
Db       2 PHQGP 6

RESULT  11
Q9X3K1
ID   Q9X3K1      PRELIMINARY;      PRT;     8 AA.
AC   Q9X3K1;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE   CYTOCHROME B (FRAGMENT).
GN   PETB.
OS   Prochlorococcus sp.
OC   Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococcaceae;
OC   Prochlorococcus.
OX   NCBI_TaxID=1220;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Urbach E., Chisholm S.W.;
RT   "Genetic diversity in Prochlorococcus populations flow cytometrically
RT   sorted from the Sargasso Sea and Gulf Stream.";
RL   Limnol. Oceanog. 43:1615-1630(1998).
DR   EMBL; AF070193; AAD23233.1; -.
FT   NON_TER       1      1
SQ   SEQUENCE   8 AA;  799 MW;  10376865B72866D3 CRC64;

  Query Match             29.3%;  Score 17;  DB 2;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 4.2e+05;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       7 GPL 9
           |||
Db       6 GPL 8

RESULT  12
Q16468
ID   Q16468      PRELIMINARY;    PRT;    8 AA.
AC   Q16468;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAY-1999 (TrEMBLrel. 10, Last annotation update)
DE   DNA FOR COSMID CC13-1134 PCR PRIMER 1 (FRAGMENT).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96435920; PubMed=8838806;
RA   James L.A., Ogilvie D.J., Yamakawa K., Nakamura Y., Stirling C.J.,
RA   Anand R.;
RT   "Walking, cloning, and mapping with YACs in 3q27: localization of five
RT   ESTs including three members of the cystatin gene family and
RT   identification of CpG islands.";
RL   Genomics 32:425-430(1996).
DR   EMBL; X88976; CAA61407.1; -.
FT   NON_TER       8      8
SQ   SEQUENCE   8 AA;  925 MW;  FD5411A7376871E6 CRC64;

  Query Match             29.3%;  Score 17;  DB 4;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 4.2e+05;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       7 GPL 9
           |||
Db       3 GPL 5

RESULT  13
Q9P0K3
ID   Q9P0K3      PRELIMINARY;      PRT;     8 AA.
AC   Q9P0K3;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   CGMP-SPECIFIC PHOSPHODIESTERASE PDE5A2 (FRAGMENT).
GN   PDE5A.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20145478; PubMed=10679249;
RA   Lin C.S., Lau A., Tu R., Lue T.F.;
RT   "Identification of three alternative first exons and an intronic
RT   promoter of human PDE5A gene.";
RL   Biochem. Biophys. Res. Commun. 268:596-602(2000).
DR   EMBL; AF155195; AAF40302.1; -.
FT   NON_TER       8      8
SQ   SEQUENCE   8 AA;  908 MW;  E8A33AA879D76726 CRC64;

  Query Match             29.3%;  Score 17;  DB 4;  Length 8;
  Best Local Similarity   50.0%;  Pred. No. 4.2e+05;
  Matches    2;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2 LMPF 5
           ::||
Db       1 MLPF 4

RESULT  14
Q9QWT0
ID   Q9QWT0      PRELIMINARY;     PRT;     9 AA.
AC   Q9QWT0;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   PROTEINASE 3 (FRAGMENT).
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SV129 D3;
RA   Lutz P.G., Moog-Lutz C., Houzel-Charavel A., Cayre Y.E.;
RL   Submitted (JUN-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AJ007030; CAA07429.1; -.
FT   NON_TER       9      9
SQ   SEQUENCE   9 AA;  937 MW;  C91E75A77B45B87D CRC64;

  Query Match             29.3%;  Score 17;  DB 11;  Length 9;
  Best Local Similarity   50.0%;  Pred. No. 4.2e+05;
  Matches    2;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       5 FPGP 8
           :| |
Db       5 YPSP 8

RESULT  15
O93739
```

```
RT   "Differential distribution of pyrokinin-isoforms in cerebral and
RT   abdominal neurohemal organs of the American cockroach.";
RL   Insect Biochem. Mol. Biol. 29:139-144(1999).
RN   [2]
RP   TISSUE SPECIFICITY.
RX   MEDLINE=20189894; PubMed=10723010;
RA   Predel R., Eckert M.;
RT   "Tagma-specific distribution of FXPRLamides in the nervous system of
RT   the American cockroach.";
RL   J. Comp. Neurol. 419:352-363(2000).
CC   -!- FUNCTION: MEDIATES VISCERAL MUSCLE CONTRACTILE ACTIVITY
CC       (MYOTROPIC ACTIVITY).
CC   -!- TISSUE SPECIFICITY: CORPORA CARDIACA.
CC   -!- MASS SPECTROMETRY: MW=996.5; METHOD=MALDI.
CC   -!- SIMILARITY: BELONGS TO THE PYROKININ FAMILY.
KW   Neuropeptide; Amidation; Pyrokinin.
FT   MOD_RES       8      8       AMIDATION.
SQ   SEQUENCE   8 AA;  997 MW;  0B34177409D772C7 CRC64;

  Query Match             31.0%;  Score 18;  DB 5;  Length 8;
  Best Local Similarity   75.0%;  Pred. No. 4.2e+05;
  Matches    3;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 LMPF 5
              |:||
Db          1 LVPF 4


RESULT   7
Q99193
ID   Q99193       PRELIMINARY;      PRT;     9 AA.
AC   Q99193;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JAN-1998 (TrEMBLrel. 05, Last annotation update)
DE   HYPOTHETICAL PROTEIN IN RPOB 3'REGION (FRAGMENT).
OS   Pseudomonas putida.
OC   Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC   Pseudomonas.
OX   NCBI_TaxID=303;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Borodin A.M., Danilkovich A.V., Allikmets R.L., Rostapshov V.M.,
RA   Chernov I.P., Azhikina T.L., Monastyrskaya S., Sverdlov D.;
RL   Dokl. Biochem. 302:1261-1265(1988).
DR   EMBL; X15849; CAA33847.1; -.
KW   Hypothetical protein.
FT   NON_TER       9      9
SQ   SEQUENCE   9 AA;  852 MW;  5B4167776DC76727 CRC64;

  Query Match             31.0%;  Score 18;  DB 2;  Length 9;
  Best Local Similarity   60.0%;  Pred. No. 4.2e+05;
  Matches    3;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 LMPFP 6
              |:| |
Db          2 LLPAP 6


RESULT   8
Q9TRU7
ID   Q9TRU7       PRELIMINARY;      PRT;     9 AA.
AC   Q9TRU7;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   GAP-3, GTPASE-ACTIVATING PROTEIN.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
```

```
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=92112868; PubMed=1309786;
RA   Nice E.C., Fabri L., Hammacher A., Holden J., Simpson R.J.,
RA   Burgess A.W.;
RT   "The purification of a Rap1 GTPase-activating protein from bovine
RT   brain cytosol.";
RL   J. Biol. Chem. 267:1546-1553(1992).
SQ   SEQUENCE   9 AA;  1063 MW;  89EDA77B47604B5A CRC64;

  Query Match             31.0%;  Score 18;  DB 6;  Length 9;
  Best Local Similarity   50.0%;  Pred. No. 4.2e+05;
  Matches    2;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 MPFP 6
              :|:|
Db          4 IPYP 7


RESULT   9
Q9QX65
ID   Q9QX65       PRELIMINARY;      PRT;     9 AA.
AC   Q9QX65;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   ZINC FINGER PROTEIN NEURO-D4 (FRAGMENT).
GN   NEURO-D4.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CD1; TISSUE=E13 TRIGEMINAL GANGLOA;
RA   Buchman V.L.;
RL   Submitted (FEB-1996) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U48239; AAF21456.1; -.
FT   NON_TER       9      9
SQ   SEQUENCE   9 AA;  928 MW;  CD0E75A76042D1AD CRC64;

  Query Match             31.0%;  Score 18;  DB 11;  Length 9;
  Best Local Similarity   75.0%;  Pred. No. 4.2e+05;
  Matches    3;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          6 PGPL 9
              | ||
Db          6 PSPL 9


RESULT  10
Q9QVE9
ID   Q9QVE9       PRELIMINARY;      PRT;    10 AA.
AC   Q9QVE9;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   PROTAMINE MP2 INTERMEDIATE PROTEIN PMP2/11.
OS   Mus sp.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10095;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=92174934; PubMed=1541289;
RA   Chauviere M., Martinage A., Debarle M., Sautiere P., Chevaillier P.;
RT   "Molecular characterization of six intermediate proteins in the
RT   processing of mouse protamine P2 precursor.";
RL . Eur. J. Biochem. 204:759-765(1992).
SQ   SEQUENCE   10 AA;  1029 MW;  FCAD2DD8676866D1 CRC64;
```

```
OC   Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC   Echinoidea; Euechinoidea; Echinacea; Temnopleuroida; Toxopneustidae;
OC   Lytechinus.
OX   NCBI_TaxID=7653;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97264487; PubMed=9110313;
RA   Cserjesi P., Fang H., Brandhorst B.P.;
RT   "Metallothionein gene expression in embryos of the sea urchin
RT   Lytechinus pictus.";
RL   Mol. Reprod. Dev. 47:39-46(1997).
DR   EMBL; U83400; AAB58320.1; -.
FT   NON_TER       8      8
SQ   SEQUENCE   8 AA;  823 MW;  EBD5A2C1F7686766 CRC64;

  Query Match             34.5%;  Score 20;  DB 5;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 4.2e+05;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       6 PGP 8
           |||
Db       2 PGP 4

RESULT   3
Q9P2Z5
ID   Q9P2Z5      PRELIMINARY;      PRT;     9 AA.
AC   Q9P2Z5;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   PML-RARA FUSION PROTEIN (FRAGMENT).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=95194921; PubMed=7534109;
RA   Yoshida H., Naoe T., Fukutani H., Kiyoi H., Kubo K., Ohno R.;
RT   "Analysis of the joining sequences of the t(15;17) translocation in
RT   human acute promyelocytic leukemia: sequence non-specific
RT   recombination between the PML and RARA genes within identical short
RT   stretches.";
RL   Genes Chromosomes Cancer 12:37-44(1995).
DR   EMBL; S76395; AAD14221.1; -.
FT   NON_TER       1      1
SQ   SEQUENCE   9 AA;  1017 MW;  5BD04362D7777767 CRC64;

  Query Match             34.5%;  Score 20;  DB 4;  Length 9;
  Best Local Similarity   50.0%;  Pred. No. 4.2e+05;
  Matches    3;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       4 PFPGPL 9
           | | |:
Db       1 PLPPPV 6

RESULT   4
Q9UNF2
ID   Q9UNF2      PRELIMINARY;      PRT;    10 AA.
AC   Q9UNF2;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   ALPHA 1 COLLAGEN (FRAGMENT).
GN   COL1A1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=BREAST;
RA   Ratcliffe T.A., Vitz J.R., Ray D.B.;
RT   "SNP located within intron 32 of human pro-alpha 1 (I) collagen gene
RT   at 10,828 bp: cytosine replaces adenine.";
RL   Submitted (FEB-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF128441; AAD32608.1; -.
KW   Collagen.
FT   NON_TER       1      1
FT   NON_TER      10     10
SQ   SEQUENCE   10 AA;  953 MW;  364F91A873276867 CRC64;

  Query Match             34.5%;  Score 20;  DB 4;  Length 10;
  Best Local Similarity   100.0%;  Pred. No. 2.6e+03;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       6 PGP 8
           |||
Db       3 PGP 5

RESULT   5
Q9UCU9
ID   Q9UCU9      PRELIMINARY;      PRT;    10 AA.
AC   Q9UCU9;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   COMPLEMENT COMPONENT C1S=SERINE PROTEASE.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=91308095; PubMed=1854725;
RA   Illy C., Thielens N.M., Gagnon J., Arlaud G.J.;
RT   "Effect of lactoperoxidase-catalyzed iodination on the Ca(2+)-
RT   dependent interactions of human C1s. Location of the iodination
RT   sites.";
RL   Biochemistry 30:7135-7141(1991).
SQ   SEQUENCE   10 AA;  1144 MW;  230113A766876AA8 CRC64;

  Query Match             32.8%;  Score 19;  DB 4;  Length 10;
  Best Local Similarity   50.0%;  Pred. No. 3.8e+03;
  Matches    4;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 HLMPFPGP 8
           |  | | |
Db       2 HGDPMPCP 9

RESULT   6
P82618
ID   P82618      PRELIMINARY;      PRT;     8 AA.
AC   P82618;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   PYROKININ-3 (PEA-PK-3) (FXPRL-AMIDE).
OS   Periplaneta americana (American cockroach).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Orthopteroidea; Dictyoptera; Blattaria;
OC   Blattoidea; Blattidae; Periplaneta.
OX   NCBI_TaxID=6978;
RN   [1]
RP   SEQUENCE, FUNCTION, AND MASS SPECTROMETRY.
RC   TISSUE=RETROCEREBRAL COMPLEX;
RX   MEDLINE=99212469; PubMed=10196736;
RA   Predel R., Kellner R., Nachman R.J., Holman G.M., Rapus J., Gaede G.;
```

```
                     GenCore version 4.5
             Copyright (c) 1993 - 2000  Compugen Ltd.


OM protein - protein search, using sw model

Run on:         October 23, 2001, 13:34:33 ; Search time 20.83 Seconds
                                             (without alignments)
                                             63.517 Million cell updates/sec

Title:          US-09-625-963-3
Perfect score:  58
Sequence:       1 HLMPFPGPLL 10

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched:       425026 seqs, 132305027 residues

Total number of hits satisfying chosen parameters:      890

Minimum DB seq length: 0
Maximum DB seq length: 10

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :      SPTREMBL_16:*
               1:  sp_archea:*
               2:  sp_bacteria:*
               3:  sp_fungi:*
               4:  sp_human:*
               5:  sp_invertebrate:*
               6:  sp_mammal:*
               7:  sp_mhc:*
               8:  sp_organelle:*
               9:  sp_phage:*
               10:  sp_plant:*
               11:  sp_rodent:*
               12:  sp_unclassified:*
               13:  sp_vertebrate:*
               14:  sp_virus:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                              SUMMARIES
                 %
Result           Query
  No.   Score    Match Length DB  ID                    Description
-------------------------------------------------------------------------------
    1      24     41.4      9  5  Q9TWV0                Q9twv0 anthopleura
    2      20     34.5      8  5  O02032                O02032 lytechinus
    3      20     34.5      9  4  Q9P2Z5                Q9p2z5 homo sapien
    4      20     34.5     10  4  Q9UNF2                Q9unf2 homo sapien
    5      19     32.8     10  4  Q9UCU9                Q9ucu9 homo sapien
    6      18     31.0      8  5  P82618                P82618 periplaneta
    7      18     31.0      9  2  Q99193                Q99193 pseudomonas
    8      18     31.0      9  6  Q9TRU7                Q9tru7 bos taurus
    9      18     31.0      9 11   Q9QX65                Q9qx65 mus musculu
   10      18     31.0     10 11   Q9QVE9                Q9qve9 mus sp. pro
   11      17     29.3      8  2  Q9X3K1                Q9x3k1 prochloroco
   12      17     29.3      8  4  Q16468                Q16468 homo sapien
   13      17     29.3      8  4  Q9P0K3                Q9p0k3 homo sapien
   14      17     29.3      9 11   Q9QWT0                Q9qwt0 mus musculu
   15      17     29.3     10  1  O93739                O93739 methanother
   16      17     29.3     10  2  Q9R5T2                Q9r5t2 acetobacter
   17      17     29.3     10 11   Q9QVF0                Q9qvf0 mus sp. pro
   18      17     29.3     10 14   Q85462                Q85462 avian sarco
   19      16     27.6      8  7  Q95213                Q95213 oryctolagus
   20      16     27.6      9  4  Q99887                Q99887 homo sapien
   21      16     27.6     10 10   P81898                P81898 prunus dulc
   22      16     27.6     10 11   Q9QVF7                Q9qvf7 rattus sp.
   23      16     27.6     10 13   P81750                P81750 clupea pall
   24      15     25.9      8  2  Q45615                Q45615 bacillus su
   25      15     25.9      8  4  Q15901                Q15901 homo sapien
   26      15     25.9      9  4  Q9UMF3                Q9umf3 homo sapien
   27      15     25.9      9  5  Q9TWD6                Q9twd6 leptinotars
   28      15     25.9      9 10   Q9S8J8                Q9s8j8 oryza sativ
   29      15     25.9      9 14   Q85599                Q85599 moloney mur
   30      15     25.9     10  4  Q9H3R9                Q9h3r9 homo sapien
   31      15     25.9     10  8  Q9MJQ5                Q9mjq5 podospora c
   32      15     25.9     10 11   Q9QV21                Q9qv21 rattus sp.
   33      15     25.9     10 13   Q9PRY8                Q9pry8 triakis scy
   34      15     25.9     10 13   Q9PRU9                Q9pru9 sparus aura
   35      15     25.9     10 14   O39957                O39957 hepatitis g
   36      15     25.9     10 14   O90346                O90346 hepatitis g
   37      14     24.1      7  8  O99182                O99182 gnatholebia
   38      14     24.1      8  6  Q9TT78                Q9tt78 canis famil
   39      14     24.1      9 14   Q69473                Q69473 human herpe
   40      14     24.1     10  5  Q9TWU1                Q9twul fusinus fer
   41      14     24.1     10 11   Q9QVJ7                Q9qvj7 mus sp. mep
   42      14     24.1     10 13   Q9PS07                Q9ps07 alligator m
   43    13.5     23.3      8  6  Q9XSY1                Q9xsyl canis famil
   44      13     22.4      7  8  O98866                O98866 spinacia ol
   45      13     22.4      7 14   Q07624                Q07624 rous sarcom


                                 ALIGNMENTS


RESULT   1
Q9TWV0
ID   Q9TWV0      PRELIMINARY;      PRT;     9 AA.
AC   Q9TWV0;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   ANTHO-RPAMIDE=NEUROPEPTIDE.
OS   Anthopleura elegantissima (Sea anemone).
OC   Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Actiniaria;
OC   Nynantheae; Actiniidae; Anthopleura.
OX   NCBI_TaxID=6110;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=93126143; PubMed=1480510;
RA   Carstensen K., Rinehart K.L., McFarlane I.D., Grimmelikhuijzen C.J.;
RT   "Isolation of Leu-Pro-Pro-Gly-Pro-Leu-Pro-Arg-Pro-NH2 (Antho-RPamide),
RT   an N-terminally protected, biologically active neuropeptide from sea
RT   anemones.";
RL   Peptides 13:851-857(1992).
SQ   SEQUENCE   9 AA;  943 MW;  2908176737686777 CRC64;


  Query Match             41.4%;  Score 24;  DB 5;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 4.2e+05;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       6 PGPL 9
           ||||
Db       3 PGPL 6


RESULT   2
O02032
ID   O02032      PRELIMINARY;      PRT;     8 AA.
AC   O02032;
DT   01-JUL-1997 (TrEMBLrel. 04, Created)
DT   01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   METALLOTHIONEIN (FRAGMENT).
GN   LPMT2.
OS   Lytechinus pictus (Painted sea urchin).
```

```
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   ANGIOTENSIN-LIKE PEPTIDE II (FRAGMENT).
OS   Bothrops jararaca (Jararaca).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC   Viperidae; Crotalinae; Bothrops.
OX   NCBI_TaxID=8724;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Plasma;
RX   MEDLINE=96208932; PubMed=8829801;
RA   Borgheresi R.A.M.B., Dalle Lucca J., Carmona E., Picarelli Z.P.;
RT   "Isolation and identification of angiotensin-like peptides from the
RT   plasma of the snake Bothrops jararaca.";
RL   Comp. Biochem. Physiol. 113B:467-473(1996).
CC   -!- SIMILARITY: BELONGS TO THE SERPIN FAMILY.
DR   InterPro; IPR000215; -.
DR   PROSITE; PS00284; SERPIN; PARTIAL.
KW   Vasoconstrictor; Plasma; Serpin.
FT   NON_TER       8      8
SQ   SEQUENCE   8 AA;  1046 MW;  DDD761E04B42D40A CRC64;


  Query Match             25.9%;  Score 15;  DB 1;  Length 8;
  Best Local Similarity   40.0%;  Pred. No. 9.3e+04;
  Matches    2;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 HLMPF 5
          :: ||
Db      4 YIHPF 8


Search completed: October 23, 2001, 13:37:42
Job time: 169 sec
```

```
RT   in avian species.";
RL   Proc. Natl. Acad. Sci. U.S.A. 81:3874-3878(1984).
RN   [2]
RP   SEQUENCE.
RC   SPECIES=A.mississippiensis; TISSUE=Brain;
RX   MEDLINE=91352338; PubMed=1882082;
RA   Lovejoy D.A., Fischer W.H., Parker D.B., McRory J.E., Park M.,
RA   Lance V., Swanson P., Rivier J.E., Sherwood N.M.;
RT   "Primary structure of two forms of gonadotropin-releasing hormone
RT   from brains of the American alligator (Alligator mississippiensis).";
RL   Regul. Pept. 33:105-116(1991).
RN   [3]
RP   SEQUENCE.
RC   SPECIES=S.acanthias; TISSUE=Brain;
RX   MEDLINE=92335300; PubMed=1631133;
RA   Lovejoy D.A., Fischer W.H., Ngamvongchon S., Craig A.G.,
RA   Nahorniak C.S., Peter R.E., Rivier J.E., Sherwood N.M.;
RT   "Distinct sequence of gonadotropin-releasing hormone (GnRH) in
RT   dogfish brain provides insight into GnRH evolution.";
RL   Proc. Natl. Acad. Sci. U.S.A. 89:6373-6377(1992).
RN   [4]
RP   SEQUENCE.
RC   SPECIES=H.colliei; TISSUE=Brain;
RX   MEDLINE=91340067; PubMed=1678723;
RA   Lovejoy D.A., Sherwood N.M., Fischer W.H., Jackson B.C., Rivier J.E.,
RA   Lee T.;
RT   "Primary structure of gonadotropin-releasing hormone from the brain
RT   of a holocephalan (ratfish: Hydrolagus colliei).";
RL   Gen. Comp. Endocrinol. 82:152-161(1991).
CC   -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.
CC   -!- SIMILARITY: BELONGS TO THE GNRH FAMILY.
DR   PIR; B60066; RHAQ2.
DR   PIR; A61126; A61126.
DR   PIR; B46030; B46030.
DR   InterPro; IPR002012; -.
DR   Pfam; PF00446; GnRH; 1.
DR   PROSITE; PS00473; GNRH; 1.
KW   Hormone; Amidation; Hypothalamus.
FT   MOD_RES       1      1       PYRROLIDONE CARBOXYLIC ACID.
FT   MOD_RES      10     10       AMIDATION.
SQ   SEQUENCE   10 AA;  1254 MW;  284B2E437871F5A3 CRC64;


  Query Match             27.6%;  Score 16;  DB 1;  Length 10;
  Best Local Similarity   66.7%;  Pred. No. 2.8e+03;
  Matches    2;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       5 FPG 7
           :||
Db       8 YPG 10


RESULT  13
RL16_ACHLA
ID   RL16_ACHLA     STANDARD;      PRT;    10 AA.
AC   P29221;
DT   01-DEC-1992 (Rel. 24, Created)
DT   01-DEC-1992 (Rel. 24, Last sequence update)
DT   01-NOV-1995 (Rel. 32, Last annotation update)
DE   50S RIBOSOMAL PROTEIN L16 (FRAGMENT).
GN   RPLP.
OS   Acholeplasma laidlawii.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC   Acholeplasmataceae; Acholeplasma.
OX   NCBI_TaxID=2148;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92210505; PubMed=1556079;
RA   Lim P.O., Sears B.B.;
RT   "Evolutionary relationships of a plant-pathogenic mycoplasmalike
RT   organism and Acholeplasma laidlawii deduced from two ribosomal protein
RT   gene sequences.";
RL   J. Bacteriol. 174:2606-2611(1992).
```

```
CC   -!- FUNCTION: THIS PROTEIN BINDS DIRECTLY TO 23S RIBOSOMAL RNA AND IS
CC       LOCATED AT THE A SITE OF THE PEPTIDYLTRANSFERASE CENTER
CC       (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE L16P FAMILY OF RIBOSOMAL PROTEINS.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; M74771; AAA21914.1; -.
DR   PIR; F41839; F41839.
DR   InterPro; IPR000114; -.
DR   PROSITE; PS00586; RIBOSOMAL_L16_1; PARTIAL.
DR   PROSITE; PS00701; RIBOSOMAL_L16_2; PARTIAL.
KW   Ribosomal protein; rRNA-binding.
FT   NON_TER      10     10
SQ   SEQUENCE   10 AA;  1324 MW;  B3386A21B4032766 CRC64;

  Query Match             27.6%;  Score 16;  DB 1;  Length 10;
  Best Local Similarity   100.0%;  Pred. No. 2.8e+03;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2 LMP 4
           |||
Db       2 LMP 4


RESULT  14
UN06_PINPS
ID   UN06_PINPS     STANDARD;      PRT;     7 AA.
AC   P81675;
DT   15-JUL-1999 (Rel. 38, Created)
DT   15-JUL-1999 (Rel. 38, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   UNKNOWN PROTEIN FROM 2D-PAGE OF NEEDLES (N141) (FRAGMENT).
OS   Pinus pinaster (Maritime pine).
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Coniferopsida; Coniferales; Pinaceae; Pinus.
OX   NCBI_TaxID=71647;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Needle;
RX   MEDLINE=99274088; PubMed=10344291;
RA   Costa P., Pionneau C., Bauw G., Dubos C., Bahrman N., Kremer A.,
RA   Frigerio J.-M., Plomion C.;
RT   "Separation and characterization of needle and xylem maritime pine
RT   proteins.";
RL   Electrophoresis 20:1098-1108(1999).
CC   -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN
CC       PROTEIN IS: 6.6, ITS MW IS: 25 KDA.
FT   NON_TER       1      1
FT   NON_TER       7      7
SQ   SEQUENCE   7 AA;  823 MW;  69D76724486B5740 CRC64;


  Query Match             25.9%;  Score 15;  DB 1;  Length 7;
  Best Local Similarity   66.7%;  Pred. No. 9.3e+04;
  Matches    2;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3 MPF 5
           :||
Db       5 LPF 7


RESULT  15
ANG2_BOTJA
ID   ANG2_BOTJA     STANDARD;      PRT;     8 AA.
AC   Q10582;
```

```
DT   01-APR-1993 (Rel. 25, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   UNKNOWN PROTEIN FROM 2D-PAGE OF PLASMA (SPOT 36) (FRAGMENT).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Plasma;
RX   MEDLINE=93092937; PubMed=1459097;
RA   Hughes G.J., Frutiger S., Paquet N., Ravier F., Pasquali C.,
RA   Sanchez J.-C., James R., Tissot J.-D., Bjellqvist B.,
RA   Hochstrasser D.F.;
RT   "Plasma protein map: an update by microsequencing.";
RL   Electrophoresis 13:707-714(1992).
CC   -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN
CC       PROTEIN IS: 7, ITS MW IS: 12 KDA.
DR   SWISS-2DPAGE; P30096; HUMAN.
FT   NON_TER       1      1
FT   VARIANT       5      5       F -> P.
FT                                /FTId=VAR_000004.
FT   NON_TER       8      8
SQ   SEQUENCE   8 AA;  909 MW;  86677B59D1A72042 CRC64;

  Query Match             27.6%;  Score 16;  DB 1;  Length 8;
  Best Local Similarity   66.7%;  Pred. No. 9.3e+04;
  Matches    2;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          5 FPG 7
              :||
Db          6 YPG 8


RESULT 10
FAR9_ASCSU
ID   FAR9_ASCSU     STANDARD;      PRT;     9 AA.
AC   P43172;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   01-FEB-1996 (Rel. 33, Last annotation update)
DE   FMRFAMIDE-LIKE NEUROPEPTIDE AF9.
OS   Ascaris suum (Pig roundworm) (Ascaris lumbricoides).
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Ascaridida; Ascaridoidea;
OC   Ascarididae; Ascaris.
OX   NCBI_TaxID=6253;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=95380362; PubMed=7651904;
RA   Cowden C., Stretton A.O.W.;
RT   "Eight novel FMRFamide-like neuropeptides isolated from the nematode
RT   Ascaris suum.";
RL   Peptides 16:491-500(1995).
CC   -!- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)
CC       FAMILY.
KW   Neuropeptide; Amidation.
FT   MOD_RES       9      9       AMIDATION.
SQ   SEQUENCE   9 AA;  1012 MW;  524F073774176877 CRC64;

  Query Match             27.6%;  Score 16;  DB 1;  Length 9;
  Best Local Similarity   75.0%;  Pred. No. 9.3e+04;
  Matches    3;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          6 PGPL 9
              | ||
Db          4 PRPL 7


RESULT 11
MGMT_BOVIN
ID   MGMT_BOVIN     STANDARD;      PRT;     9 AA.
AC   P29177;
DT   01-DEC-1992 (Rel. 24, Created)
DT   01-DEC-1992 (Rel. 24, Last sequence update)
DT   01-OCT-1996 (Rel. 34, Last annotation update)
DE   METHYLATED-DNA--PROTEIN-CYSTEINE METHYLTRANSFERASE (EC 2.1.1.63) (6-O-
DE   METHYLGUANINE-DNA METHYLTRANSFERASE) (FRAGMENT).
GN   MGMT.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Thymus;
RX   MEDLINE=90174912; PubMed=2308822;
RA   Rydberg B., Hall J., Karran P.;
RT   "Active site amino acid sequence of the bovine O6-methylguanine-DNA
RT   methyltransferase.";
RL   Nucleic Acids Res. 18:17-21(1990).
CC   -!- FUNCTION: REPAIR OF ALKYLATED GUANINE IN DNA BY STOICHIOMETRICALLY
CC       TRANSFERRING THE ALKYL GROUP AT THE O-6 POSITION TO A CYSTEINE
CC       RESIDUE IN THE ENZYME. THIS IS A SUICIDE REACTION: THE ENZYME IS
CC       IRREVERSIBLY INACTIVATED.
CC   -!- CATALYTIC ACTIVITY: DNA (CONTAINING O6-METHYLGUANINE) + PROTEIN
CC       L-CYSTEINE = DNA (WITHOUT O6-METHYLGUANINE) + PROTEIN S-METHYL-
CC       L-CYSTEINE.
CC   -!- SIMILARITY: WITH SEGMENTS OF E.COLI ADA AND OGT METHYLTRANSFERASE
CC       WHICH ENCOMPASS THE ALKYL-ACCEPTOR RESIDUES.
DR   InterPro; IPR001497; -.
DR   PROSITE; PS00374; MGMT; PARTIAL.
KW   DNA repair; Transferase; Methyltransferase.
FT   NON_TER       1      1
FT   ACT_SITE      9      9       ALKYL GROUP ACCEPTOR (BY SIMILARITY).
FT   NON_TER       9      9
SQ   SEQUENCE   9 AA;  967 MW;  325171A720476047 CRC64;

  Query Match             27.6%;  Score 16;  DB 1;  Length 9;
  Best Local Similarity   60.0%;  Pred. No. 9.3e+04;
  Matches    3;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          6 PGPLL 10
              | |:|
Db          2 PIPIL 6


RESULT 12
GON2_CHICK
ID   GON2_CHICK     STANDARD;      PRT;    10 AA.
AC   P37043; P20408;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   GONADOLIBERIN II (GONADOTROPIN-RELEASING HORMONE II) (GNRH-II)
DE   (LH-RH II) (LULIBERIN II).
OS   Gallus gallus (Chicken),
OS   Alligator mississippiensis (American alligator),
OS   Squalus acanthias (Spiny dogfish), and
OS   Hydrolagus colliei (Spotted ratfish) (Pacific ratfish).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031, 8496, 7797, 7873;
RN   [1]
RP   SEQUENCE.
RC   SPECIES=Chicken; TISSUE=Hypothalamus;
RX   MEDLINE=84222059; PubMed=6427779;
RA   Miyamoto K., Hasegawa Y., Nomura M., Igarashi M., Kangawa K.,
RA   Matsuo H.;
RT   "Identification of the second gonadotropin-releasing hormone in
RT   chicken hypothalamus: evidence that gonadotropin secretion is
RT   probably controlled by two distinct gonadotropin-releasing hormones
```

```
AC   P80465;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   01-FEB-1996 (Rel. 33, Last annotation update)
DE   QUINOLINE 2-OXIDOREDUCTASE, BETA CHAIN (EC 1.-.-.-) (FRAGMENT).
OS   Comamonas testosteroni (Pseudomonas testosteroni).
OC   Bacteria; Proteobacteria; beta subdivision; Comamonadaceae; Comamonas.
OX   NCBI_TaxID=285;
RN   [1]
RP   SEQUENCE.
RC   STRAIN=63;
RX   MEDLINE=96035889; PubMed=7556204;
RA   Schach S., Tshisuaka B., Fetzner S., Lingens F.;
RT   "Quinoline 2-oxidoreductase and 2-oxo-1,2-dihydroquinoline 5,6-
RT   dioxygenase from Comamonas testosteroni 63. The first two enzymes in
RT   quinoline and 3-methylquinoline degradation.";
RL   Eur. J. Biochem. 232:536-544(1995).
CC   -!- FUNCTION: CONVERTS (3-METHYL-)-QUINOLINE TO (3-METHYL-)2-OXO-
CC       1,2-DIHYDROQUINOLINE.
CC   -!- COFACTOR: FAD AND MOLYBDENUM.
CC   -!- PATHWAY: FIRST STEP IN THE DEGRADATION OF QUINOLINE AND
CC       (3-METHYL-)QUINOLINE.
CC   -!- SUBUNIT: HETEROHEXAMER OF TWO ALPHA CHAINS, TWO BETA CHAINS, AND
CC       TWO GAMMA CHAINS (PROBABLE).
KW   Oxidoreductase; Flavoprotein; FAD; Molybdenum.
FT   NON_TER      10     10
SQ   SEQUENCE   10 AA;  1241 MW;  C2E2C25DD9CDC769 CRC64;


  Query Match             29.3%;  Score 17;  DB 1;  Length 10;
  Best Local Similarity   75.0%;  Pred. No. 1.9e+03;
  Matches    3;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       3 MPFP 6
           | ||
Db       1 MKFP 4


RESULT   7
TMOF_AEDAE
ID   TMOF_AEDAE     STANDARD;      PRT;    10 AA.
AC   P19425;
DT   01-NOV-1990 (Rel. 16, Created)
DT   01-NOV-1990 (Rel. 16, Last sequence update)
DT   01-FEB-1996 (Rel. 33, Last annotation update)
DE   TRYPSIN-MODULATING OOSTATIC FACTOR (TMOF) (OOSH).
OS   Aedes aegypti (Yellowfever mosquito).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea;
OC   Culicidae; Aedes.
OX   NCBI_TaxID=7159;
RN   [1]
RP   SEQUENCE.
RC   STRAIN=VERO BEACH; TISSUE=Ovary;
RX   MEDLINE=90367888; PubMed=2394318;
RA   Borovsky D., Carlson D.A., Griffin P.R., Shabanowitz J., Hunt D.F.;
RT   "Mosquito oostatic factor: a novel decapeptide modulating
RT   trypsin-like enzyme biosynthesis in the midgut.";
RL   FASEB J. 4:3015-3020(1990).
RN   [2]
RP   SEQUENCE.
RC   STRAIN=VERO BEACH; TISSUE=Ovary;
RX   MEDLINE=93357794; PubMed=8353526;
RA   Borovsky D., Carlson D.A., Griffin P.R., Shabanowitz J., Hunt D.F.;
RT   "Mass spectrometry and characterization of Aedes aegypti trypsin
RT   modulating oostatic factor (TMOF) and its analogs.";
RL   Insect Biochem. Mol. Biol. 23:703-712(1993).
CC   -!- FUNCTION: HAS AN OOSTATIC ACTIVITY. INHIBITS TRYPSIN BIOSYNTHESIS
CC       IN THE MIDGUT WHICH INDIRECTLY REDUCES THE VITELLOGENIN
CC       CONCENTRATION IN THE HEMOLYMPH RESULTING IN INHIBITION OF OOCYTE
CC       DEVELOPMENT.
CC   -!- DEVELOPMENTAL STAGE: SYNTHESIZED AND RELEASED FROM FOLLICULAR
CC       EPITHELIUM 18-24 HRS AFTER A BLOOD MEAL. SYNTHESIS PEAKS AT
CC       36 HRS AND STOPS AT 56 HRS.
DR   PIR; A36454; A36454.
KW   Hormone.
FT   DOMAIN        3     10       POLY-PRO.
FT   VARIANT       1      2       YD -> DY (IN TMFO(B)).
SQ   SEQUENCE   10 AA;  1047 MW;  236D0A7777776DC7 CRC64;


  Query Match             29.3%;  Score 17;  DB 1;  Length 10;
  Best Local Similarity   60.0%;  Pred. No. 1.9e+03;
  Matches    3;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       4 PFPGP 8
           | | |
Db       3 PAPPP 7


RESULT   8
RS7_MYCIT
ID   RS7_MYCIT      STANDARD;      PRT;     8 AA.
AC   P33564;
DT   01-FEB-1994 (Rel. 28, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   01-FEB-1994 (Rel. 28, Last annotation update)
DE   30S RIBOSOMAL PROTEIN S7 (FRAGMENT).
GN   RPSG.
OS   Mycobacterium intracellulare.
OC   Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC   Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX   NCBI_TaxID=1767;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93197130; PubMed=8451173;
RA   Nair J., Rouse D.A., Morris S.L.;
RT   "Nucleotide sequence analysis of the ribosomal S12 gene of
RT   Mycobacterium intracellulare.";
RL   Nucleic Acids Res. 21:1039-1039(1993).
CC   -!- FUNCTION: PROTEIN S7 BINDS SPECIFICALLY TO PART OF THE 3' END OF
CC       16S RIBOSOMAL RNA (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE S7P FAMILY OF RIBOSOMAL PROTEINS.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; L08171; AAA25376.1; -.
DR   PIR; S35538; S35538.
DR   InterPro; IPR000235; -.
DR   PROSITE; PS00052; RIBOSOMAL_S7; PARTIAL.
KW   Ribosomal protein; rRNA-binding.
FT   INIT_MET      0      0       BY SIMILARITY.
FT   NON_TER       8      8
SQ   SEQUENCE   8 AA;  850 MW;  63276DC768732417 CRC64;


  Query Match             27.6%;  Score 16;  DB 1;  Length 8;
  Best Local Similarity   60.0%;  Pred. No. 9.3e+04;
  Matches    3;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       4 PFPGP 8
           |  ||
Db       1 PRKGP 5


RESULT   9
UPAA_HUMAN
ID   UPAA_HUMAN     STANDARD;      PRT;     8 AA.
AC   P30096;
DT   01-APR-1993 (Rel. 25, Created)
```

```
CC   -!- FUNCTION: THIS PEPTIDE BOTH INHIBITS THE ACTIVITY OF THE
CC       ANGIOTENSIN-CONVERTING ENZYME AND ENHANCES THE ACTION OF
CC       BRADYKININ BY INHIBITING THE KINASES THAT INACTIVATE IT.
CC       IT ACTS AS AN INDIRECT HYPOTENSIVE AGENT.
DR   PIR; A60377; XASNPC.
KW   Hypotensive agent; Venom.
FT   MOD_RES       1      1       PYRROLIDONE CARBOXYLIC ACID.
SQ   SEQUENCE   10 AA;  1062 MW;  3BA827C327686773 CRC64;


  Query Match             36.2%;  Score 21;  DB 1;  Length 10;
  Best Local Similarity   75.0%;  Pred. No. 4.1e+02;
  Matches    3;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        5 FPGP 8
            :|||
Db        3 WPGP 6


RESULT   3
SAMP_MUSCA
ID   SAMP_MUSCA     STANDARD;      PRT;     9 AA.
AC   P19095;
DT   01-NOV-1990 (Rel. 16, Created)
DT   01-NOV-1990 (Rel. 16, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   SERUM AMYLOID P-COMPONENT (SAP) (FRAGMENT).
OS   Mustelus canis (Smooth dogfish).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC   Elasmobranchii; Galeomorphii; Galeoidea; Carcharhiniformes; Triakidae;
OC   Mustelus.
OX   NCBI_TaxID=7812;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=83160932; PubMed=6403520;
RA   Robey F.A., Tanaka T., Liu T.-Y.;
RT   "Isolation and characterization of two major serum proteins from the
RT   dogfish, Mustelus canis, C-reactive protein and amyloid P
RT   component.";
RL   J. Biol. Chem. 258:3889-3894(1983).
CC   -!- SUBUNIT: HOMOPENTAMER. PENTAXIN (OR PENTRAXIN) HAVE A DISCOID
CC       ARRANGEMENT OF 5 NONCOVALENTLY BOUND SUBUNITS.
CC   -!- DISEASE: SAP IS A PRECURSOR OF AMYLOID COMPONENT P WHICH IS FOUND
CC       IN BASEMENT MEMBRANE AND ASSOCIATED WITH AMYLOID DEPOSITS.
CC   -!- SIMILARITY: BELONGS TO THE PENTAXIN FAMILY.
DR   PIR; B20569; B20569.
DR   InterPro; IPR001759; -.
DR   PROSITE; PS00289; PENTAXIN; PARTIAL.
KW   Amyloid; Glycoprotein; Plasma; Pentaxin.
FT   DOMAIN        1     >9       PENTAXIN.
FT   NON_TER       9      9
SQ   SEQUENCE   9 AA;  965 MW;  D05B5735B3386769 CRC64;


  Query Match             34.5%;  Score 20;  DB 1;  Length 9;
  Best Local Similarity   66.7%;  Pred. No. 9.3e+04;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy        5 FPGPLL 10
            |||  |
Db        2 FPGKSL 7


RESULT   4
MNP1_LEPDE
ID   MNP1_LEPDE     STANDARD;      PRT;     7 AA.
AC   P42984;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   01-NOV-1995 (Rel. 32, Last annotation update)
DE   MYOTROPIC NEUROPEPTIDE 1 (LED-MNP-I).
OS   Leptinotarsa decemlineata (Colorado potato beetle).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;
OC   Cucujiformia; Phytophaga; Chrysomeloidea; Chrysomelidae;
OC   Chrysomelinae; Leptinotarsa.
OX   NCBI_TaxID=7539;
RN   [1]
RP   SEQUENCE, AND SYNTHESIS.
RC   TISSUE=Head;
RX   MEDLINE=95380343; PubMed=7651886;
RA   Spittaels K., Vankeerberghen A., Schoofs L., Torrekens S.,
RA   Grauwels L., van Leuven F., de Loof A.;
RT   "Identification, characterization, and immunological localization of
RT   a novel myotropic neuropeptide in the Colorado potato beetle,
RT   Leptinotarsa decemlineata.";
RL   Peptides 16:365-374(1995).
CC   -!- FUNCTION: MYOTROPIC PEPTIDE. STIMULATES THE CONTRACTIONS OF THE
CC       OVIDUCT.
KW   Neuropeptide; Amidation.
FT   MOD_RES       7      7       AMIDATION.
SQ   SEQUENCE   7 AA;  705 MW;  6DD73768745B5DB0 CRC64;


  Query Match             31.0%;  Score 18;  DB 1;  Length 7;
  Best Local Similarity   60.0%;  Pred. No. 9.3e+04;
  Matches    3;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        5 FPGPL 9
             : |||
Db        2 YNGPL 6


RESULT   5
GRP_RANRI
ID   GRP_RANRI      STANDARD;      PRT;    10 AA.
AC   P23260;
DT   01-NOV-1991 (Rel. 20, Created)
DT   01-NOV-1991 (Rel. 20, Last sequence update)
DT   01-APR-1993 (Rel. 25, Last annotation update)
DE   NEUROMEDIN C.
OS   Rana ridibunda (Laughing frog) (Marsh frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Neobatrachia; Ranoidea; Ranidae; Rana.
OX   NCBI_TaxID=8406;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Brain;
RX   MEDLINE=91315477; PubMed=1859413;
RA   Conlon J.M., O'Harte F., Vaudry H.;
RT   "Primary structures of the bombesin-like neuropeptides in frog brain
RT   show that bombesin is not the amphibian gastrin-releasing peptide.";
RL   Biochem. Biophys. Res. Commun. 178:526-530(1991).
CC   -!- SIMILARITY: BELONGS TO THE BOMBESIN/NEUROMEDIN B/RANATENSIN
CC       FAMILY.
DR   PIR; PQ0177; PQ0177.
DR   InterPro; IPR000874; -.
DR   Pfam; PF02044; Bombesin; 1.
DR   PROSITE; PS00257; BOMBESIN; 1.
KW   Bombesin family; Amidation.
FT   MOD_RES      10     10       AMIDATION.
SQ   SEQUENCE   10 AA;  1094 MW;  F81FBAE862CDC371 CRC64;


  Query Match             29.3%;  Score 17;  DB 1;  Length 10;
  Best Local Similarity   100.0%;  Pred. No. 1.9e+03;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 HLM 3
            |||
Db        8 HLM 10


RESULT   6
Q2OB_COMTE
ID   Q2OB_COMTE     STANDARD;      PRT;    10 AA.
```

```
                    GenCore version 4.5
             Copyright (c) 1993 - 2000  Compugen Ltd.


OM protein - protein search, using sw model

Run on:          October 23, 2001, 13:34:53 ; Search time 9.99 Seconds
                                               (without alignments)
                                               34.290 Million cell updates/sec

Title:           US-09-625-963-3
Perfect score:   58
Sequence:        1 HLMPFPGPLL 10

Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched:        93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters:     331

Minimum DB seq length: 0
Maximum DB seq length: 10

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       SwissProt_39:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                                SUMMARIES

                  %
Result          Query
  No.   Score   Match  Length DB  ID                    Description
-------------------------------------------------------------------------
    1     22    37.9      8  1  VGLG_HSV2B             P81780 herpes simp
    2     21    36.2     10  1  BPP_VIPAS              P31351 vipera aspi
    3     20    34.5      9  1  SAMP_MUSCA             P19095 mustelus ca
    4     18    31.0      7  1  MNP1_LEPDE             P42984 leptinotars
    5     17    29.3     10  1  GRP_RANRI              P23260 rana ridibu
    6     17    29.3     10  1  Q2OB_COMTE             P80465 comamonas t
    7     17    29.3     10  1  TMOF_AEDAE             P19425 aedes aegyp
    8     16    27.6      8  1  RS7_MYCIT              P33564 mycobacteri
    9     16    27.6      8  1  UPAA_HUMAN             P30096 homo sapien
   10     16    27.6      9  1  FAR9_ASCSU             P43172 ascaris suu
   11     16    27.6      9  1  MGMT_BOVIN             P29177 bos taurus
   12     16    27.6     10  1  GON2_CHICK             P37043 gallus gall
   13     16    27.6     10  1  RL16_ACHLA             P29221 acholeplasm
   14     15    25.9      7  1  UNO6_PINPS             P81675 pinus pinas
   15     15    25.9      8  1  ANG2_BOTJA             Q10582 bothrops ja
   16     15    25.9     10  1  COXO_THUOB             P80982 thunnus obe
   17     14    24.1     10  1  ANG1_BOTJA             Q10581 bothrops ja
   18     14    24.1     10  1  ANGT_BOVIN             P01017 bos taurus
   19     14    24.1     10  1  ANGT_CHICK             P01018 gallus gall
   20     14    24.1     10  1  COXO_RAT               P80432 rattus norv
   21     14    24.1     10  1  TRP5_LEUMA             P81737 leucophaea
   22     14    24.1     10  1  TRP9_LEUMA             P81741 leucophaea
   23     13    22.4      5  1  PAP2_PARMA             P81864 pardachirus
   24     13    22.4      5  1  SUGA_ACHDO             P19991 acheta dome
   25     13    22.4      6  1  TRPI_PSEPU             P36414 pseudomonas
   26     13    22.4      7  1  CARP_MYTED             P10420 mytilus edu
   27     13    22.4      7  1  FAR1_HELTI             P41871 helisoma tr
   28     13    22.4      8  1  AKH_TABAT              P14595 tabanus atr
   29     13    22.4      8  1  AL15_CARMA             P81818 carcinus ma
   30     13    22.4      8  1  AL16_CARMA             P81819 carcinus ma
   31     13    22.4      8  1  ALL5_CALVO             P41841 calliphora
   32     13    22.4      8  1  ALL8_CARMA             P81811 carcinus ma
   33     13    22.4      8  1  ALL9_CARMA             P81812 carcinus ma
```

```
   34     13    22.4      8  1  FAR7_ASCSU             P43171 ascaris suu
   35     13    22.4      8  1  FUSS_FUSSO             P81010 fusarium so
   36     13    22.4      8  1  LCK2_LEUMA             P21141 leucophaea
   37     13    22.4      8  1  RPCH_PANBO             P08939 pandalus bo
   38     13    22.4      8  1  UF06_MOUSE             P38644 mus musculu
   39     13    22.4      9  1  CCAP_CARMA             P38556 carcinus ma
   40     13    22.4      9  1  FAR5_CALVO             P41860 calliphora
   41     13    22.4      9  1  LITO_LITAU             P08945 litoria aur
   42     13    22.4      9  1  LITR_PHYRO             P08946 phyllomedus
   43     13    22.4      9  1  LMT3_LOCMI             P41489 locusta mig
   44     13    22.4      9  1  MOSH_CLYJA             P19852 clypeaster
   45     13    22.4      9  1  TKL1_LOCMI             P16223 locusta mig


                                ALIGNMENTS


RESULT   1
VGLG_HSV2B
ID   VGLG_HSV2B     STANDARD;      PRT;     8 AA.
AC   P81780;
DT   15-JUL-1999 (Rel. 38, Created)
DT   15-JUL-1999 (Rel. 38, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   GLYCOPROTEIN G (FRAGMENT).
OS   Herpes simplex virus (type 2 / strain B4327UR).
OC   Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC   Alphaherpesvirinae; Simplexvirus.
OX   NCBI_TaxID=103921;
RN   [1]
RP   SEQUENCE.
RA   Liljeqvist J.-A., Svennerholm B., Bergstrom T.;
RL   Submitted (APR-1999) to the SWISS-PROT data bank.
CC   -!- MISCELLANEOUS: THERE ARE SEVEN EXTERNAL GLYCOPROTEINS IN HSV1 AND
CC       2: GH, GB, GC, GG, GD, GI, AND GE.
CC   -!- MISCELLANEOUS: GLYCOPROTEIN G IS MUCH LARGER IN HSV-2 THAN IN
CC       HSV-1.
KW   Glycoprotein.
FT   NON_TER       8      8
SQ   SEQUENCE   8 AA;  683 MW;  7B47686772C865B8 CRC64;


  Query Match             37.9%;  Score 22;  DB 1;  Length 8;
  Best Local Similarity   75.0%;  Pred. No. 9.3e+04;
  Matches    3;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        6 PGPL 9
            |||:
Db        5 PGPI 8


RESULT   2
BPP_VIPAS
ID   BPP_VIPAS      STANDARD;      PRT;    10 AA.
AC   P31351;
DT   01-JUL-1993 (Rel. 26, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   01-FEB-1994 (Rel. 28, Last annotation update)
DE   BRADYKININ-POTENTIATING PEPTIDE (ANGIOTENSIN-CONVERTING
DE   ENZYME INHIBITOR).
OS   Vipera aspis (Aspic viper).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC   Viperidae; Viperinae; Vipera.
OX   NCBI_TaxID=8706;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Venom;
RX   MEDLINE=90382616; PubMed=2169439;
RA   Komori Y., Sugihara H.;
RT   "Characterization of a new inhibitor for angiotensin converting
RT   enzyme from the venom of Vipera aspis aspis.";
RL   Int. J. Biochem. 22:767-771(1990).
```

```
neuromedin C - laughing frog
C;Species: Rana ridibunda (laughing frog)
C;Date: 23-Nov-1991 #sequence_revision 23-Nov-1991 #text_change 11-Jan-2000
C;Accession: PQ0177
R;Conlon, J.M.; O'Harte, F.; Vaudry, H.
Biochem. Biophys. Res. Commun. 178, 526-530, 1991
A;Title: Primary structures of the bombesin-like neuropeptides in frog brain show that b
A;Reference number: PQ0177; MUID:91315477
A;Accession: PQ0177
A;Molecule type: protein
A;Residues: 1-10 <CON>
A;Experimental source: brain
C;Superfamily: gastrin-releasing peptide
C;Keywords: amidated carboxyl end
F;10/Modified site: amidated carboxyl end (Met) #status predicted


  Query Match             29.3%;  Score 17;  DB 2;  Length 10;
  Best Local Similarity   100.0%;  Pred. No. 3.5e+03;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 HLM 3
            |||
Db        8 HLM 10

RESULT  15
A60647
neuromedin C - bovine
C;Species: Bos primigenius taurus (cattle)
C;Date: 14-May-1993 #sequence_revision 14-May-1993 #text_change 07-May-1999
C;Accession: A60647
R;Lemaire, S.; Trifaro, J.M.; Chouinard, L.; Cecyre, D.; Dessureault, J.; Mercier, P[illegible]D
Peptides 10, 355-360, 1989
A;Title: Structural identification, subcellular localization and secretion of bovine adr
A;Reference number: A60647; MUID:89331342
A;Accession: A60647
A;Molecule type: protein
A;Residues: 1-10 <LEM>
A;Note: this neuropeptide was purified from secretory granules of cells in the adrenal m
C;Superfamily: gastrin-releasing peptide
C;Keywords: adrenal gland; neuropeptide


  Query Match             29.3%;  Score 17;  DB 2;  Length 10;
  Best Local Similarity   100.0%;  Pred. No. 3.5e+03;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 HLM 3
            |||
Db        8 HLM 10


Search completed: October 23, 2001, 13:34:48
Job time: 260 sec
```

```
A;Accession: C39111
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-10 <VAR>
C;Keywords: heterotetramer; immunoglobulin


  Query Match             34.5%;  Score 20;  DB 2;  Length 10;
  Best Local Similarity   44.4%;  Pred. No. 1.1e+03;
  Matches    4;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          2 LMPFPGPLL 10
              | |   ||:
Db          1 LNPISSPLV 9


RESULT   9
A46306
spasmogenic toxin PNV1 - spider (Phoneutria nigriventer) (fragment)
C;Species: Phoneutria nigriventer
C;Date: 07-Oct-1994 #sequence_revision 07-Oct-1994 #text_change 17-Mar-1999
C;Accession: A46306
R;Marangoni, S.; Borges, N.C.C.; Marangoni, R.A.; Antunes, E.; Vieira, C.A.; Novello, J.
Toxicon 31, 377-384, 1993
A;Title: Biochemical characterization of a vascular smooth muscle contracting polypeptid
A;Reference number: A46306; MUID:93276438
A;Accession: A46306
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-8 <MAR>


  Query Match             32.8%;  Score 19;  DB 2;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 2.2e+05;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          5 FPG 7
              |||
Db          3 FPG 5


RESULT  10
S10783
enamelin f - bovine (fragment)
C;Species: Bos primigenius taurus (cattle)
C;Date: 19-Mar-1997 #sequence_revision 21-Nov-1998 #text_change 21-Nov-1998
C;Accession: S10783
R;Strawich, E.; Glimcher, M.J.
Eur. J. Biochem. 191, 47-56, 1990
A;Title: Tooth 'enamelins' identified mainly as serum proteins. Major 'enamelin' is albu
A;Reference number: S10780; MUID:90336641
A;Accession: S10783
A;Molecule type: protein
A;Residues: 1-8 <STR>
C;Keywords: enamel; phosphoprotein


  Query Match             32.8%;  Score 19;  DB 2;  Length 8;
  Best Local Similarity   75.0%;  Pred. No. 2.2e+05;
  Matches    3;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          3 MPFP 6
              || |
Db          1 MPLP 4


RESULT  11
S71948
matrix metalloproteinase 3 precursor - bovine (fragments)
N;Alternate names: MMP-3 protein; stromelysin precursor MMP-3
C;Species: Bos primigenius taurus (cattle)
C;Date: 19-Mar-1998 #sequence_revision 17-Apr-1998 #text_change 07-May-1999
C;Accession: S71948
```

```
R;Arner, E.C.; Pratta, M.A.; Freimark, B.; Lischwe, M.; Trzaskos, J.M.; Magolda, R.L.
Biochem. J. 318, 417-424, 1996
A;Title: Isothiazolones interfere with normal matrix metalloproteinase activation and
A;Reference number: S71948; MUID:96404887
A;Accession: S71948
A;Molecule type: protein
A;Residues: 1-5;6-10 <ARN>
C;Function:
A;Description: degrades a wide range of extracellular matrix components, including ca
A;Note: may be involved in arthritis formation


  Query Match             32.8%;  Score 19;  DB 2;  Length 10;
  Best Local Similarity   100.0%;  Pred. No. 1.6e+03;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          5 FPG 7
              |||
Db          8 FPG 10


RESULT  12
H28027
protein P11 - curled-leaved tobacco (fragment)
C;Species: Nicotiana plumbaginifolia (curled-leaved tobacco)
C;Date: 19-May-1989 #sequence_revision 19-May-1989 #text_change 18-Jun-1993
C;Accession: H28027
R;Bauw, G.; De Loose, M.; Inze, D.; Van Montagu, M.; Vandekerckhove, J.
Proc. Natl. Acad. Sci. U.S.A. 84, 4806-4810, 1987
A;Title: Alterations in the phenotype of plant cells studied by NH2-terminal amino ac
A;Reference number: A94167
A;Accession: H28027
A;Molecule type: protein
A;Residues: 1-10 <BAU>
A;Note: 4-Val was also found


  Query Match             31.0%;  Score 18;  DB 2;  Length 10;
  Best Local Similarity   75.0%;  Pred. No. 2.4e+03;
  Matches    3;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          5 FPGP 8
              | ||
Db          4 FEGP 7


RESULT  13
S66607
quinoline 2-oxidoreductase beta chain - Comamonas testosteroni (fragment)
C;Species: Comamonas testosteroni
C;Date: 15-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 17-Mar-1999
C;Accession: S66607
R;Schach, S.; Tshisuaka, B.; Fetzner, S.; Lingens, F.
Eur. J. Biochem. 232, 536-544, 1995
A;Title: Quinoline 2-oxidoreductase and 2-oxo-1,2-dihydroquinoline 5,6-dioxygenase fr
A;Reference number: S66606; MUID:96035889
A;Accession: S66607
A;Molecule type: protein
A;Residues: 1-9 <SCH>
A;Experimental source: strain 63


  Query Match             29.3%;  Score 17;  DB 2;  Length 9;
  Best Local Similarity   75.0%;  Pred. No. 2.2e+05;
  Matches    3;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          3 MPFP 6
              | ||
Db          1 MKFP 4


RESULT  14
PQ0177
```

```
RESULT   3
D46285
formaldehyde dehydrogenase (glutathione) (EC 1.2.1.1) class III low activity form - Atla
C;Species: Gadus morhua (Atlantic cod)
C;Date: 19-Nov-1993 #sequence_revision 18-Nov-1994 #text_change 20-Mar-1996
C;Accession: D46285
R;Danielsson, O.; Jornvall, H.
Proc. Natl. Acad. Sci. U.S.A. 89, 9247-9251, 1992
A;Title: "Enzymogenesis": classical liver alcohol dehydrogenase origin from the glutathi
A;Reference number: A46285; MUID:93028441
A;Accession: D46285
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-10 <DAN>
A;Note: sequence extracted from NCBI backbone (NCBIP:116272)
C;Keywords: NAD; oxidoreductase


  Query Match             37.9%;  Score 22;  DB 2;  Length 10;
  Best Local Similarity   60.0%;  Pred. No. 4.9e+02;
  Matches    3;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 HLMPF 5
           | :||
Db       6 HTLPF 10


RESULT   4
S35538
ribosomal protein S7 - Mycobacterium intracellulare (fragment)
C;Species: Mycobacterium intracellulare
C;Date: 09-Dec-1993 #sequence_revision 14-Nov-1997 #text_change 13-Aug-1999
C;Accession: S35538
R;Nair, J.; Rouse, D.; Morris, S.
Nucleic Acids Res. 21, 1039, 1993
A;Title: Nucleotide sequence analysis of the ribosomal S12 gene of Mycobacterium intrace
A;Reference number: S35537; MUID:93197130
A;Accession: S35538
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-9 <NAI>
A;Cross-references: EMBL:L08171; NID:g149994; PIDN:AAA25376.1; PID:g551901
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, February 1993
C;Superfamily: Escherichia coli ribosomal protein S7
C;Keywords: protein biosynthesis; ribosome


  Query Match             36.2%;  Score 21;  DB 2;  Length 9;
  Best Local Similarity   66.7%;  Pred. No. 2.2e+05;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       3 MPFPGP 8
           ||  ||
Db       1 MPRKGP 6


RESULT   5
S26508
collagen alpha 2(VI) chain - bovine (fragment)
C;Species: Bos primigenius taurus (cattle)
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 05-Dec-1998
C;Accession: S26508
R;Jander, R.; Rautenberg, J.; Glanville, R.W.
Eur. J. Biochem. 133, 39-46, 1983
A;Title: Further characterization of the three polypeptide chains of bovine and human sh
A;Reference number: S26506; MUID:83209648
A;Accession: S26508
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-9 <JAN>
C;Keywords: hydroxyproline
F;7/Modified site: hydroxyproline (Pro) #status experimental
```

```
  Query Match             36.2%;  Score 21;  DB 2;  Length 9;
  Best Local Similarity   50.0%;  Pred. No. 2.2e+05;
  Matches    3;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       3 MPFPGP 8
           :  |||
Db       1 LEIPGP 6


RESULT   6
XASNPC
angiotensin-converting enzyme inhibitor - aspic viper
C;Species: Vipera aspis (aspic viper)
C;Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 08-Dec-1995
C;Accession: A60377
R;Komori, Y.; Sugihara, H.
Int. J. Biochem. 22, 767-771, 1990
A;Title: Characterization of a new inhibitor for angiotensin converting enzyme from t
A;Reference number: A60377; MUID:90382616
A;Accession: A60377
A;Molecule type: protein
A;Residues: 1-10 <KOM>
C;Superfamily: bradykinin-potentiating peptide
C;Keywords: angiotensin-converting enzyme inhibitor; pyroglutamic acid
F;1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental


  Query Match             36.2%;  Score 21;  DB 1;  Length 10;
  Best Local Similarity   75.0%;  Pred. No. 7.3e+02;
  Matches    3;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       5 FPGP 8
           :|||
Db       3 WPGP 6


RESULT   7
B20569
serum amyloid P-component - smooth dogfish (fragment)
C;Species: Mustelus canis (smooth dogfish)
C;Date: 05-Jun-1987 #sequence_revision 05-Jun-1987 #text_change 18-Jun-1993
C;Accession: B20569; A05074
R;Robey, F.A.; Tanaka, T.; Liu, T.Y.
J. Biol. Chem. 258, 3889-3894, 1983
A;Title: Isolation and characterization of two major serum proteins from the dogfish,
A;Reference number: A92419; MUID:83160932
A;Accession: B20569
A;Molecule type: protein
A;Residues: 1-9 <ROB>
C;Keywords: amyloid


  Query Match             34.5%;  Score 20;  DB 2;  Length 9;
  Best Local Similarity   66.7%;  Pred. No. 2.2e+05;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       5 FPGPLL 10
           |||  |
Db       2 FPGKSL 7


RESULT   8
C39111
Ig heavy chain C region - Pacific hagfish (fragment)
C;Species: Eptatretus stouti (Pacific hagfish)
C;Date: 18-Oct-1991 #sequence_revision 18-Oct-1991 #text_change 16-Aug-1996
C;Accession: C39111
R;Varner, J.; Neame, P.; Litman, G.W.
Proc. Natl. Acad. Sci. U.S.A. 88, 1746-1750, 1991
A;Title: A serum heterodimer from hagfish (Eptatretus stoutii) exhibits structural si
A;Reference number: A39111; MUID:91156684
```

```
                    GenCore version 4.5
             Copyright (c) 1993 - 2000  Compugen Ltd.


OM protein - protein search, using sw model

Run on:          October 23, 2001, 13:30:28 ; Search time 12.71 Seconds
                                              (without alignments)
                                              59.933 Million cell updates/sec

Title:           US-09-625-963-3
Perfect score:   58
Sequence:        1 HLMPFPGPLL 10

Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched:        219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters:      1098

Minimum DB seq length: 0
Maximum DB seq length: 10

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       PIR_68:*
                1:  pir1:*
                2:  pir2:*
                3:  pir3:*
                4:  pir4:*

     Pred. No. is the number of results predicted by chance to have a
     score greater than or equal to the score of the result being printed,
     and is derived by analysis of the total score distribution.

                                   SUMMARIES

                      %
Result              Query
  No.    Score   Match Length DB  ID                        Description
 -------------------------------------------------------------------------------
    1       35    60.3      6  2  A61049                    halo-toxin - Pseud
    2       23    39.7     10  2  S68033                    cytochrome P450 1A
    3       22    37.9     10  2  D46285                    formaldehyde dehyd
    4       21    36.2      9  2  S35538                    ribosomal protein
    5       21    36.2      9  2  S26508                    collagen alpha 2(V
    6       21    36.2     10  1  XASNPC                    angiotensin-conver
    7       20    34.5      9  2  B20569                    serum amyloid P-co
    8       20    34.5     10  2  C39111                    Ig heavy chain C r
    9       19    32.8      8  2  A46306                    spasmogenic toxin
   10       19    32.8      8  2  S10783                    enamelin f - bovin
   11       19    32.8     10  2  S71948                    matrix metalloprot
   12       18    31.0     10  2  H28027                    protein P11 - curl
   13       17    29.3      9  2  S66607                    quinoline 2-oxidor
   14       17    29.3     10  2  PQ0177                    neuromedin C - lau
   15       17    29.3     10  2  A60647                    neuromedin C - bov
   16       17    29.3     10  2  S18396                    probable glucose-6
   17       17    29.3     10  2  A36454                    trypsin-modulating
   18       16    27.6      4  2  PT0240                    Ig heavy chain CRD
   19       16    27.6      7  2  E33932                    Ig mu chain D regi
   20       16    27.6     10  1  RHAQ2                     gonadoliberin II -
   21       16    27.6     10  1  A61126                    gonadoliberin - sp
   22       16    27.6     10  2  F41839                    ribosomal protein
   23       16    27.6     10  2  B59272                    peptide-N4-(N-acet
   24       16    27.6     10  2  PQ0783                    NADH dehydrogenase
   25       16    27.6     10  2  B46030                    gonadoliberin II -
   26       16    27.6     10  2  S26506                    collagen alpha 1(V
   27       15    25.9      5  2  S53595                    hypothetical prote
   28       15    25.9      7  2  S66442                    glutathione S-tran
   29       15    25.9      8  2  S66646                    cardioacceleratory
   30       15    25.9      8  2  PT0368                    Ig gamma chain C r
   31       15    25.9      8  2  A35180                    neutral proteinase
   32       15    25.9      9  2  S10784                    enamelin i - bovin
   33       15    25.9     10  2  C39745                    sphingomyelinase -
   34       15    25.9     10  2  PX0030                    triacylglycerol li
   35       15    25.9     10  2  S65432                    angiotensin I - ho
   36       15    25.9     10  2  S77990                    cytochrome-c oxida
   37       14    24.1      5  2  B60274                    major protein anti
   38       14    24.1      7  2  B48394                    major fat-globule
   39       14    24.1      8  2  S21288                    lectin - potato (f
   40       14    24.1      9  2  B45020                    probable minipolyp
   41       14    24.1      9  2  A60108                    exotoxin A - Strep
   42       14    24.1     10  2  S65388                    cytochrome-c oxida
   43       14    24.1     10  2  A60624                    angiotensin I - Ja
   44       14    24.1     10  2  A39745                    endo-glucosylceram
   45       14    24.1     10  2  A90917                    angiotensin precur


                                  ALIGNMENTS


RESULT   1
A61049
halo-toxin - Pseudomonas syringae pv. mori
C;Species: Pseudomonas syringae pv. mori
A;Note: host mulberry tree
C;Date: 10-Mar-1994 #sequence_revision 10-Mar-1994 #text_change 21-Jan-1997
C;Accession: A61049
R;Kajimoto, T.; Yokomizo, K.; Yahiro, K.; Umeda, T.; Shoji, S.; Kubota, Y.; Shibata,
Chem. Lett. 00, 679-680, 1989
A;Title: Structure of halo-toxin produced by phytopathogenic bacterium, Pseudomonas s
A;Reference number: A61049
A;Accession: A61049
A;Molecule type: protein
A;Residues: 1-6 <KAJ>
A;Note: sequence confirmed by synthesis
C;Comment: This toxin is one of the etiological agents of halo bright disease in mulb
C;Keywords: toxin


  Query Match             60.3%;  Score 35;  DB 2;  Length 6;
  Best Local Similarity   83.3%;  Pred. No. 2.2e+05;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        4 PFPGPL 9
              |||||:
Db        1 PFPGPI 6


RESULT   2
S68033
cytochrome P450 1A1 - tilapia (fragment)
C;Species: Oreochromis niloticus x Oreochromis aureus (tilapia)
C;Date: 19-Mar-1997 #sequence_revision 29-Aug-1997 #text_change 29-Aug-1997
C;Accession: S68033
R;Ueng, Y.F.; Ueng, T.H.
Arch. Biochem. Biophys. 322, 347-356, 1995
A;Title: Induction and purification of cytochrome P450 1A1 from 3-methylcholanthrene-
A;Reference number: S68033; MUID:96032654
A;Accession: S68033
A;Molecule type: protein
A;Residues: 1-10 <UEN>
A;Experimental source: liver and gill


  Query Match             39.7%;  Score 23;  DB 2;  Length 10;
  Best Local Similarity   50.0%;  Pred. No. 3.3e+02;
  Matches    4;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy        2 LMPFPGPL 9
              ::|| | |
Db        1 ILPFIGAL 8
```

```
;       STREET:  One Liberty Place - 46th Floor
;       CITY:  Philadelphia
;       STATE:  PA
;       COUNTRY:  USA
;       ZIP:  19103
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  DISKETTE, 3.5 INCH, 1.44 Mb STORAGE
;       COMPUTER:  IBM PS/2
;       OPERATING SYSTEM:  PC-DOS
;       SOFTWARE:  WORDPERFECT 5.0
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  PCT/US93/08436
;       FILING DATE:  Herewith
;       CLASSIFICATION:
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:
;       FILING DATE:
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Dianne B. Elderkin
;       REGISTRATION NUMBER:  28,598
;       REFERENCE/DOCKET NUMBER:  CCOR-0023
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (215) 568-3100
;       TELEFAX:  (215) 568-3439
;   INFORMATION FOR SEQ ID NO:  15:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  9 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:
;       TOPOLOGY:  unknown
;     MOLECULE TYPE:
;       DESCRIPTION:  Cyclic structure involving cyclization
;       DESCRIPTION:  between Cys(2) and Cys(5);  Amide terminated.
PCT-US93-08436-15


  Query Match            42.9%;  Score 24;  DB 5;  Length 9;
  Best Local Similarity   75.0%;  Pred. No. 1.5e+05;
  Matches    3;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 CMTW 4
           | ||
Db       5 CWTW 8


RESULT  15
US-08-439-817-190
; Sequence 190, Application US/08439817
; Patent No. 5728802
;  GENERAL INFORMATION:
;     APPLICANT:  Barrett, Ronald W.
;     APPLICANT:  Cwirla, Steven E.
;     APPLICANT:  Dower, William J.
;     APPLICANT:  Koller, Kerry J.
;     APPLICANT:  Lee, Jung
;     APPLICANT:  Martens, Christine L.
;     APPLICANT:  Ruhland-Fritsch, Beatrice
;     TITLE OF INVENTION:  Peptides and Compounds That Bind
;     TITLE OF INVENTION:  Selectins Including Endothelial Leukocyte Adhesion
;     TITLE OF INVENTION:  Molecule I (ELAM-1)
;     NUMBER OF SEQUENCES:  209
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Affymax Technologies, NV
;       STREET:  4001 Miranda Ave.
;       CITY:  Palo Alto
;       STATE:  California
;       COUNTRY:  USA
;       ZIP:  94304
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
```

```
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/439,817
;       FILING DATE:  12-MAY-1995
;       CLASSIFICATION:  530
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/241,054
;       FILING DATE:  11-MAY-1994
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/057,295
;       FILING DATE:  05-MAY-1993
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 07/881,395
;       FILING DATE:  06-MAY-1992
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Stevens, Lauren L.
;       REGISTRATION NUMBER:  36,691
;       REFERENCE/DOCKET NUMBER:  000324-046/1056.1
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  415-496-2300
;       TELEFAX:  415-424-0832
;   INFORMATION FOR SEQ ID NO:  190:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  5 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
;     FEATURE:
;       NAME/KEY:  Region
;       LOCATION:  one-of(1)
;       OTHER INFORMATION:  /note= "N-terminal Ile is
;       OTHER INFORMATION:  benzyloxycarboyl."
;     FEATURE:
;       NAME/KEY:  Region
;       LOCATION:  one-of(5)
;       OTHER INFORMATION:  /note= "C-terminal Gln is amidated."
US-08-439-817-190


  Query Match            41.1%;  Score 23;  DB 1;  Length 5;
  Best Local Similarity   60.0%;  Pred. No. 1.5e+05;
  Matches    3;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2 MTWNQ 6
           :||:|
Db       1 ITWDQ 5


Search completed: October 23, 2001, 13:29:22
Job time: 390 sec
```

```
;  INFORMATION FOR SEQ ID NO:  15:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      OTHER INFORMATION:  Cyclic structure involving
;      OTHER INFORMATION:  cyclization between Cys(2) and Cys(5);  Amide
;      OTHER INFORMATION:  terminated.
US-08-397-101-15


  Query Match             42.9%;  Score 24;  DB 1;  Length 9;
  Best Local Similarity   75.0%;  Pred. No. 1.5e+05;
  Matches    3;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 CMTW 4
           | ||
Db       5 CWTW 8


RESULT  12
PCT-US93-08436-9
; Sequence 9, Application PC/TUS9308436
;  GENERAL INFORMATION:
;    APPLICANT:  George A. Heavner, et al.
;    TITLE OF INVENTION:  Peptide Inhibitors of Cellular
;    TITLE OF INVENTION:  Adhesion
;    NUMBER OF SEQUENCES:  19
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Woodcock Washburn Kurtz Mackiewicz &
;      ADDRESSEE:  Norris
;      STREET:  One Liberty Place - 46th Floor
;      CITY:  Philadelphia
;      STATE:  PA
;      COUNTRY:  USA
;      ZIP:  19103
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  DISKETTE, 3.5 INCH, 1.44 Mb STORAGE
;      COMPUTER:  IBM PS/2
;      OPERATING SYSTEM:  PC-DOS
;      SOFTWARE:  WORDPERFECT 5.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US93/08436
;      FILING DATE:  Herewith
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Dianne B. Elderkin
;      REGISTRATION NUMBER:  28,598
;      REFERENCE/DOCKET NUMBER:  CCOR-0023
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (215) 568-3100
;      TELEFAX:  (215) 568-3439
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:
;      DESCRIPTION:  Cyclic structure involving cyclization
;      DESCRIPTION:  between Cys(1) and Cys(5);  Amide terminated.
PCT-US93-08436-9


  Query Match             42.9%;  Score 24;  DB 5;  Length 9;
  Best Local Similarity   75.0%;  Pred. No. 1.5e+05;
  Matches    3;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 CMTW 4
           | ||
Db       5 CWTW 8


RESULT  13
PCT-US93-08436-10
; Sequence 10, Application PC/TUS9308436
;  GENERAL INFORMATION:
;    APPLICANT:  George A. Heavner, et al.
;    TITLE OF INVENTION:  Peptide Inhibitors of Cellular
;    TITLE OF INVENTION:  Adhesion
;    NUMBER OF SEQUENCES:  19
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Woodcock Washburn Kurtz Mackiewicz &
;      ADDRESSEE:  Norris
;      STREET:  One Liberty Place - 46th Floor
;      CITY:  Philadelphia
;      STATE:  PA
;      COUNTRY:  USA
;      ZIP:  19103
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  DISKETTE, 3.5 INCH, 1.44 Mb STORAGE
;      COMPUTER:  IBM PS/2
;      OPERATING SYSTEM:  PC-DOS
;      SOFTWARE:  WORDPERFECT 5.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US93/08436
;      FILING DATE:  Herewith
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Dianne B. Elderkin
;      REGISTRATION NUMBER:  28,598
;      REFERENCE/DOCKET NUMBER:  CCOR-0023
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (215) 568-3100
;      TELEFAX:  (215) 568-3439
;  INFORMATION FOR SEQ ID NO:  10:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:
;      DESCRIPTION:  Cyclic structure involving cyclization
;      DESCRIPTION:  between Cys(3) and Cys(6);  Amide terminated.
PCT-US93-08436-10


  Query Match             42.9%;  Score 24;  DB 5;  Length 9;
  Best Local Similarity   75.0%;  Pred. No. 1.5e+05;
  Matches    3;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 CMTW 4
           | ||
Db       6 CWTW 9


RESULT  14
PCT-US93-08436-15
; Sequence 15, Application PC/TUS9308436
;  GENERAL INFORMATION:
;    APPLICANT:  George A. Heavner, et al.
;    TITLE OF INVENTION:  Peptide Inhibitors of Cellular
;    TITLE OF INVENTION:  Adhesion
;    NUMBER OF SEQUENCES:  19
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Woodcock Washburn Kurtz Mackiewicz &
;      ADDRESSEE:  Norris
```

```
;      APPLICATION NUMBER:  US/08/397,101
;      FILING DATE:  07-MAR-1995
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/941,653
;      FILING DATE:  08-SEP-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US93/08504
;      FILING DATE:  08-SEP-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Dianne B. Elderkin
;      REGISTRATION NUMBER:  28,598
;      REFERENCE/DOCKET NUMBER:  CCOR-0139
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (215) 568-3100
;      TELEFAX:  (215) 568-3439
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      OTHER INFORMATION:  Cyclic structure involving
;      OTHER INFORMATION:  cyclization between Cys(1) and Cys(5);  Amide
;      OTHER INFORMATION:  terminated.
US-08-397-101-9

  Query Match             42.9%;  Score 24;  DB 1;  Length 9;
  Best Local Similarity   75.0%;  Pred. No. 1.5e+05;
  Matches    3;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 CMTW 4
           | ||
Db       5 CWTW 8


RESULT  10
US-08-397-101-10
; Sequence 10, Application US/08397101
; Patent No. 5753617
;  GENERAL INFORMATION:
;    APPLICANT:  Heavner, George A.
;    APPLICANT:  Kruszynski, Marian
;    APPLICANT:  Falcone, Margaret L.
;    TITLE OF INVENTION:  Peptide Inhibitors of
;    TITLE OF INVENTION:  Cellular Adhesion
;    NUMBER OF SEQUENCES:  19
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Woodcock Washburn Kurtz
;      ADDRESSEE:  Mackiewicz & No. 5753617ris
;      STREET:  One Liberty Place - 46th Floor
;      CITY:  Philadelphia
;      STATE:  PA
;      COUNTRY:  USA
;      ZIP:  19103
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  DISKETTE, 3.5 INCH, 1.44 Mb
;      MEDIUM TYPE:  STORAGE
;      COMPUTER:  IBM PC Compatible
;      OPERATING SYSTEM:  PC-DOS
;      SOFTWARE:  WORDPERFECT 5.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/397,101
;      FILING DATE:  07-MAR-1995
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/941,653
;      FILING DATE:  08-SEP-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US93/08504
```

```
;      FILING DATE:  08-SEP-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Dianne B. Elderkin
;      REGISTRATION NUMBER:  28,598
;      REFERENCE/DOCKET NUMBER:  CCOR-0139
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (215) 568-3100
;      TELEFAX:  (215) 568-3439
;  INFORMATION FOR SEQ ID NO:  10:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      OTHER INFORMATION:  Cyclic structure involving
;      OTHER INFORMATION:  cyclization between Cys(3) and Cys(6);  Amide
;      OTHER INFORMATION:  terminated.
US-08-397-101-10

  Query Match             42.9%;  Score 24;  DB 1;  Length 9;
  Best Local Similarity   75.0%;  Pred. No. 1.5e+05;
  Matches    3;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 CMTW 4
           | ||
Db       6 CWTW 9


RESULT  11
US-08-397-101-15
; Sequence 15, Application US/08397101
; Patent No. 5753617
;  GENERAL INFORMATION:
;    APPLICANT:  Heavner, George A.
;    APPLICANT:  Kruszynski, Marian
;    APPLICANT:  Falcone, Margaret L.
;    TITLE OF INVENTION:  Peptide Inhibitors of
;    TITLE OF INVENTION:  Cellular Adhesion
;    NUMBER OF SEQUENCES:  19
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Woodcock Washburn Kurtz
;      ADDRESSEE:  Mackiewicz & No. 5753617ris
;      STREET:  One Liberty Place - 46th Floor
;      CITY:  Philadelphia
;      STATE:  PA
;      COUNTRY:  USA
;      ZIP:  19103
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  DISKETTE, 3.5 INCH, 1.44 Mb
;      MEDIUM TYPE:  STORAGE
;      COMPUTER:  IBM PC Compatible
;      OPERATING SYSTEM:  PC-DOS
;      SOFTWARE:  WORDPERFECT 5.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/397,101
;      FILING DATE:  07-MAR-1995
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/941,653
;      FILING DATE:  08-SEP-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US93/08504
;      FILING DATE:  08-SEP-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Dianne B. Elderkin
;      REGISTRATION NUMBER:  28,598
;      REFERENCE/DOCKET NUMBER:  CCOR-0139
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (215) 568-3100
;      TELEFAX:  (215) 568-3439
```

```
US-08-397-101-16
; Sequence 16, Application US/08397101
; Patent No. 5753617
;  GENERAL INFORMATION:
;    APPLICANT:  Heavner, George A.
;    APPLICANT:  Kruszynski, Marian
;    APPLICANT:  Falcone, Margaret L.
;    TITLE OF INVENTION:  Peptide Inhibitors of
;    TITLE OF INVENTION:  Cellular Adhesion
;    NUMBER OF SEQUENCES:  19
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Woodcock Washburn Kurtz
;      ADDRESSEE:  Mackiewicz & No. 5753617ris
;      STREET:  One Liberty Place - 46th Floor
;      CITY:  Philadelphia
;      STATE:  PA
;      COUNTRY:  USA
;      ZIP:  19103
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  DISKETTE, 3.5 INCH, 1.44 Mb
;      MEDIUM TYPE:  STORAGE
;      COMPUTER:  IBM PC Compatible
;      OPERATING SYSTEM:  PC-DOS
;      SOFTWARE:  WORDPERFECT 5.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/397,101
;      FILING DATE:  07-MAR-1995
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/941,653
;      FILING DATE:  08-SEP-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US93/08504
;      FILING DATE:  08-SEP-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Dianne B. Elderkin
;      REGISTRATION NUMBER:  28,598
;      REFERENCE/DOCKET NUMBER:  CCOR-0139
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (215) 568-3100
;      TELEFAX:  (215) 568-3439
;  INFORMATION FOR SEQ ID NO:  16:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  8 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      OTHER INFORMATION:  Cyclic structure involving
;      OTHER INFORMATION:  cyclization between Cys(1) and Cys(4);  Amide
;      OTHER INFORMATION:  terminated.
US-08-397-101-16


  Query Match             42.9%;  Score 24;  DB 1;  Length 8;
  Best Local Similarity   75.0%;  Pred. No. 1.5e+05;
  Matches    3;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 CMTW 4
          | ||
Db      4 CWTW 7

RESULT   8
PCT-US93-08436-16
; Sequence 16, Application PC/TUS9308436
;  GENERAL INFORMATION:
;    APPLICANT:  George A. Heavner, et al.
;    TITLE OF INVENTION:  Peptide Inhibitors of Cellular
;    TITLE OF INVENTION:  Adhesion
;    NUMBER OF SEQUENCES:  19
;    CORRESPONDENCE ADDRESS:
```

```
;      ADDRESSEE:  Woodcock Washburn Kurtz Mackiewicz &
;      ADDRESSEE:  Norris
;      STREET:  One Liberty Place - 46th Floor
;      CITY:  Philadelphia
;      STATE:  PA
;      COUNTRY:  USA
;      ZIP:  19103
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  DISKETTE, 3.5 INCH, 1.44 Mb STORAGE
;      COMPUTER:  IBM PS/2
;      OPERATING SYSTEM:  PC-DOS
;      SOFTWARE:  WORDPERFECT 5.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US93/08436
;      FILING DATE:  Herewith
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Dianne B. Elderkin
;      REGISTRATION NUMBER:  28,598
;      REFERENCE/DOCKET NUMBER:  CCOR-0023
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (215) 568-3100
;      TELEFAX:  (215) 568-3439
;  INFORMATION FOR SEQ ID NO:  16:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  8 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:
;      DESCRIPTION:  Cyclic structure involving cyclization
;      DESCRIPTION:  between Cys(1) and Cys(4);  Amide terminated.
PCT-US93-08436-16


  Query Match             42.9%;  Score 24;  DB 5;  Length 8;
  Best Local Similarity   75.0%;  Pred. No. 1.5e+05;
  Matches    3;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 CMTW 4
          | ||
Db      4 CWTW 7


RESULT   9
US-08-397-101-9
; Sequence 9, Application US/08397101
; Patent No. 5753617
;  GENERAL INFORMATION:
;    APPLICANT:  Heavner, George A.
;    APPLICANT:  Kruszynski, Marian
;    APPLICANT:  Falcone, Margaret L.
;    TITLE OF INVENTION:  Peptide Inhibitors of
;    TITLE OF INVENTION:  Cellular Adhesion
;    NUMBER OF SEQUENCES:  19
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Woodcock Washburn Kurtz
;      ADDRESSEE:  Mackiewicz & No. 5753617ris
;      STREET:  One Liberty Place - 46th Floor
;      CITY:  Philadelphia
;      STATE:  PA
;      COUNTRY:  USA
;      ZIP:  19103
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  DISKETTE, 3.5 INCH, 1.44 Mb
;      MEDIUM TYPE:  STORAGE
;      COMPUTER:  IBM PC Compatible
;      OPERATING SYSTEM:  PC-DOS
;      SOFTWARE:  WORDPERFECT 5.0
;    CURRENT APPLICATION DATA:
```

```
;    APPLICANT:  Dower, William J.
;    APPLICANT:  Koller, Kerry J.
;    APPLICANT:  Lee, Jung
;    APPLICANT:  Martens, Christine L.
;    APPLICANT:  Ruhland-Fritsch, Beatrice
;    TITLE OF INVENTION:  Peptides and Compounds That Bind
;    TITLE OF INVENTION:  Selectins Including Endothelial Leukocyte Adhesion
;    TITLE OF INVENTION:  Molecule I (ELAM-1)
;    NUMBER OF SEQUENCES:  209
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Affymax Technologies, NV
;      STREET:  4001 Miranda Ave.
;      CITY:  Palo Alto
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/439,817
;      FILING DATE:  12-MAY-1995
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/241,054
;      FILING DATE:  11-MAY-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/057,295
;      FILING DATE:  05-MAY-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/881,395
;      FILING DATE:  06-MAY-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Stevens, Lauren L.
;      REGISTRATION NUMBER:  36,691
;      REFERENCE/DOCKET NUMBER:  000324-046/1056.1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-496-2300
;      TELEFAX:  415-424-0832
;  INFORMATION FOR SEQ ID NO:  194:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  6 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  Region
;      LOCATION:  one-of(1)
;      OTHER INFORMATION:  /note= "N-terminal Ile is
;      OTHER INFORMATION:  benzyloxycarboyl."
;    FEATURE:
;      NAME/KEY:  Region
;      LOCATION:  one-of(6)
;      OTHER INFORMATION:  /note= "C-terminal Leu is amidated."
US-08-439-817-194


  Query Match             44.6%;  Score 25;  DB 1;  Length 6;
  Best Local Similarity   50.0%;  Pred. No. 1.5e+05;
  Matches    3;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        2 MTWNQM 7
            :||:|:
Db        1 ITWDQL 6


RESULT   6
US-08-439-817-191
; Sequence 191, Application US/08439817
; Patent No. 5728802
```

```
;  GENERAL INFORMATION:
;    APPLICANT:  Barrett, Ronald W.
;    APPLICANT:  Cwirla, Steven E.
;    APPLICANT:  Dower, William J.
;    APPLICANT:  Koller, Kerry J.
;    APPLICANT:  Lee, Jung
;    APPLICANT:  Martens, Christine L.
;    APPLICANT:  Ruhland-Fritsch, Beatrice
;    TITLE OF INVENTION:  Peptides and Compounds That Bind
;    TITLE OF INVENTION:  Selectins Including Endothelial Leukocyte Adhesion
;    TITLE OF INVENTION:  Molecule I (ELAM-1)
;    NUMBER OF SEQUENCES:  209
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Affymax Technologies, NV
;      STREET:  4001 Miranda Ave.
;      CITY:  Palo Alto
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/439,817
;      FILING DATE:  12-MAY-1995
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/241,054
;      FILING DATE:  11-MAY-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/057,295
;      FILING DATE:  05-MAY-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/881,395
;      FILING DATE:  06-MAY-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Stevens, Lauren L.
;      REGISTRATION NUMBER:  36,691
;      REFERENCE/DOCKET NUMBER:  000324-046/1056.1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-496-2300
;      TELEFAX:  415-424-0832
;  INFORMATION FOR SEQ ID NO:  191:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  6 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  Region
;      LOCATION:  one-of(1)
;      OTHER INFORMATION:  /note= "N-terminal Thr is
;      OTHER INFORMATION:  benzyloxycarboyl."
;    FEATURE:
;      NAME/KEY:  Region
;      LOCATION:  one-of(6)
;      OTHER INFORMATION:  /note= "C-terminal Trp is amidated."
US-08-439-817-191


  Query Match             42.9%;  Score 24;  DB 1;  Length 6;
  Best Local Similarity   60.0%;  Pred. No. 1.5e+05;
  Matches    3;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        3 TWNQM 7
            ||:|:
Db        1 TWDQL 5


RESULT   7
```

```
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-042-107-64


  Query Match             50.0%;  Score 28;  DB 4;  Length 9;
  Best Local Similarity   37.5%;  Pred. No. 1.5e+05;
  Matches    3;  Conservative    4;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 CMTWNQMN 8
           ||:|: ::
Db       1 CMSWDAVS 8


RESULT   3
US-08-487-006-147
; Sequence 147, Application US/08487006
; Patent No. 5641861
;  GENERAL INFORMATION:
;    APPLICANT:  Dooley, Colette T.
;    APPLICANT:  Houghten, Richard A.
;    TITLE OF INVENTION:  No. 5641861el Mu Opioid Receptor Ligands:
;    TITLE OF INVENTION:  Agonists and Antagonists
;    NUMBER OF SEQUENCES:  222
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Campbell and Flores
;      STREET:  4370 La Jolla Village Drive, Suite 700
;      CITY:  San Diego
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  92122
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/487,006
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Campbell, Cathryn A.
;      REGISTRATION NUMBER:  31,815
;      REFERENCE/DOCKET NUMBER:  P-TP 1706
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619) 535-9001
;      TELEFAX:  (619) 535-8949
;  INFORMATION FOR SEQ ID NO:  147:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  6 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    FEATURE:
;      NAME/KEY:  Peptide
;      LOCATION:  1..6
;      OTHER INFORMATION:  /note= "With the exception of Gly
;      OTHER INFORMATION:  at position 5, all the amino acids are the D-amino
;      OTHER INFORMATION:  acids."
;    FEATURE:
;      NAME/KEY:  Peptide
;      LOCATION:  6
;      OTHER INFORMATION:  /note= "Amino acid is amidated at
;      OTHER INFORMATION:  the C-terminal."
US-08-487-006-147


  Query Match             46.4%;  Score 26;  DB 1;  Length 6;
  Best Local Similarity   80.0%;  Pred. No. 1.5e+05;
  Matches    4;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 MTWNQ 6
           ||| |
Db       2 MTWGQ 6


RESULT   4
US-08-488-659A-147
; Sequence 147, Application US/08488659A
; Patent No. 5919897
;  GENERAL INFORMATION:
;    APPLICANT:  Dooley, Colette T.
;    APPLICANT:  Houghten, Richard A.
;    TITLE OF INVENTION:  Mu Opioid Receptor Ligands:
;    TITLE OF INVENTION:  Agonists and Antagonists
;    NUMBER OF SEQUENCES:  223
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Campbell and Flores
;      STREET:  4370 La Jolla Village Drive, Suite 700
;      CITY:  San Diego
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  92122
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/488,659A
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Campbell, Cathryn A.
;      REGISTRATION NUMBER:  31,815
;      REFERENCE/DOCKET NUMBER:  P-TP 1705
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619) 535-9001
;      TELEFAX:  (619) 535-8949
;  INFORMATION FOR SEQ ID NO:  147:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  6 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    FEATURE:
;      NAME/KEY:  Peptide
;      LOCATION:  1..6
;      OTHER INFORMATION:  /note= "With the exception of Gly
;      OTHER INFORMATION:  at position 5, all the amino acids are the D-amino
;      OTHER INFORMATION:  acids."
;    FEATURE:
;      NAME/KEY:  Peptide
;      LOCATION:  6
;      OTHER INFORMATION:  /note= "Amino acid is amidated at
;      OTHER INFORMATION:  the C-terminal."
US-08-488-659A-147


  Query Match             46.4%;  Score 26;  DB 2;  Length 6;
  Best Local Similarity   80.0%;  Pred. No. 1.5e+05;
  Matches    4;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 MTWNQ 6
           ||| |
Db       2 MTWGQ 6


RESULT   5
US-08-439-817-194
; Sequence 194, Application US/08439817
; Patent No. 5728802
;  GENERAL INFORMATION:
;    APPLICANT:  Barrett, Ronald W.
;    APPLICANT:  Cwirla, Steven E.
```

```
                    GenCore version 4.5
          Copyright (c) 1993 - 2000  Compugen Ltd.


OM protein - protein search, using sw model

Run on:         October 23, 2001, 13:29:21 ; Search time 18.81 Seconds
                                              (without alignments)
                                              9.852 Million cell updates/sec

Title:          US-09-625-963-2
Perfect score:  56
Sequence:       1 CMTWNQMNL 9

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched:       197339 seqs, 20590346 residues

Total number of hits satisfying chosen parameters:      50818

Minimum DB seq length: 0
Maximum DB seq length: 9

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       Issued_Patents_AA:*
                1:  /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
                2:  /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
                3:  /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
                4:  /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
                5:  /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
                6:  /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

     Pred. No. is the number of results predicted by chance to have a
     score greater than or equal to the score of the result being printed,
     and is derived by analysis of the total score distribution.

                                 SUMMARIES
               %
Result       Query
  No.  Score Match Length DB  ID                    Description
-------------------------------------------------------------------------
    1     28  50.0      9  4  US-09-258-754-64        Sequence 64, Appl
    2     28  50.0      9  4  US-09-042-107-64        Sequence 64, Appl
    3     26  46.4      6  1  US-08-487-006-147       Sequence 147, App
    4     26  46.4      6  2  US-08-488-659A-147      Sequence 147, App
    5     25  44.6      6  1  US-08-439-817-194       Sequence 194, App
    6     24  42.9      6  1  US-08-439-817-191       Sequence 191, App
    7     24  42.9      8  1  US-08-397-101-16        Sequence 16, Appl
    8     24  42.9      8  5  PCT-US93-08436-16       Sequence 16, Appl
    9     24  42.9      9  1  US-08-397-101-9         Sequence 9, Appli
   10     24  42.9      9  1  US-08-397-101-10        Sequence 10, Appl
   11     24  42.9      9  1  US-08-397-101-15        Sequence 15, Appl
   12     24  42.9      9  5  PCT-US93-08436-9        Sequence 9, Appli
   13     24  42.9      9  5  PCT-US93-08436-10       Sequence 10, Appl
   14     24  42.9      9  5  PCT-US93-08436-15       Sequence 15, Appl
   15     23  41.1      5  1  US-08-439-817-190       Sequence 190, App
   16     23  41.1      5  1  US-08-439-817-195       Sequence 195, App
   17     23  41.1      5  1  US-08-439-817-197       Sequence 197, App
   18     23  41.1      5  1  US-08-439-817-198       Sequence 198, App
   19     23  41.1      5  1  US-08-439-817-199       Sequence 199, App
   20     23  41.1      6  1  US-08-375-911A-4        Sequence 4, Appli
   21     23  41.1      6  1  US-08-487-006-149       Sequence 149, App
   22     23  41.1      6  2  US-08-488-659A-149      Sequence 149, App
   23     23  41.1      7  1  US-08-299-567-2         Sequence 2, Appli
   24     23  41.1      8  1  US-08-526-710-8         Sequence 8, Appli
   25     23  41.1      8  3  US-08-862-855-8         Sequence 8, Appli
   26     23  41.1      8  4  US-08-444-818-445       Sequence 445, App
   27     23  41.1      8  4  US-08-444-818-446       Sequence 446, App
   28     23  41.1      8  4  US-08-444-818-447       Sequence 447, App
   29     23  41.1      9  1  US-07-646-531D-10       Sequence 10, Appl
   30     23  41.1      9  2  US-08-488-273-10        Sequence 10, Appl
   31     23  41.1      9  2  US-08-417-174-75        Sequence 75, Appl
   32     23  41.1      9  3  US-08-159-339A-980      Sequence 980, App
   33     23  41.1      9  6  5426100-10             Patent No. 5426100
   34     23  41.1      9  6  5426100-11             Patent No. 5426100
   35     22  39.3      6  1  US-08-487-006-143       Sequence 143, App
   36     22  39.3      6  1  US-08-487-006-144       Sequence 144, App
   37     22  39.3      6  2  US-08-488-659A-143      Sequence 143, App
   38     22  39.3      6  2  US-08-488-659A-144      Sequence 144, App
   39     22  39.3      8  4  US-08-444-818-448       Sequence 448, App
   40     22  39.3      9  4  US-09-258-754-339       Sequence 339, App
   41     22  39.3      9  4  US-09-042-107-339       Sequence 339, App
   42     21  37.5      4  6  5171684-24             Patent No. 5171684
   43     21  37.5      6  1  US-08-487-006-7         Sequence 7, Appli
   44     21  37.5      6  1  US-08-487-006-130       Sequence 130, App
   45     21  37.5      6  1  US-08-487-006-131       Sequence 131, App


                                ALIGNMENTS


RESULT   1
US-09-258-754-64
; Sequence 64, Application US/09258754
; Patent No. 6174687
; GENERAL INFORMATION:
;  APPLICANT: Ruoslahti, Erkki
;  APPLICANT: Pasqualini, Renata
;  APPLICANT: Rajotte, Daniel
;  TITLE OF INVENTION: Methods of Identifying Lung Homing Molecules Using
;  TITLE OF INVENTION: Membrane Dipeptidase
;  FILE REFERENCE: P-LJ 3443
;  CURRENT APPLICATION NUMBER: US/09/258,754
;  CURRENT FILING DATE: 1999-02-26
;  EARLIER APPLICATION NUMBER: 09/042,107
;  EARLIER FILING DATE: 1998-03-13
;  NUMBER OF SEQ ID NOS: 452
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 64
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-258-754-64


  Query Match             50.0%;  Score 28;  DB 4;  Length 9;
  Best Local Similarity   37.5%;  Pred. No. 1.5e+05;
  Matches    3;  Conservative    4;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        1 CMTWNQMN 8
            ||:|: ::
Db        1 CMSWDAVS 8


RESULT   2
US-09-042-107-64
; Sequence 64, Application US/09042107
; Patent No. 6232287
; GENERAL INFORMATION:
;  APPLICANT: Ruoslahti, Erkki
;  APPLICANT: Pasqualini, Renata
;  TITLE OF INVENTION: Molecules that Home to Various Selected Organs or
;  TITLE OF INVENTION: Tissues
;  FILE REFERENCE: P-LJ 2892
;  CURRENT APPLICATION NUMBER: US/09/042,107
;  CURRENT FILING DATE: 1998-03-13
;  NUMBER OF SEQ ID NOS: 436
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 64
```

```
;  APPLICANT:  Keogh, Elissa
;  APPLICANT:  Epimmune Inc.
;  TITLE OF INVENTION: Inducing Cellular Immune Responses to
;  TITLE OF INVENTION:  Prostate Cancer Antigens Using Peptide and Nucleic Acid
;  TITLE OF INVENTION:  Compositions
;  FILE REFERENCE: 018623-014710US
;  CURRENT APPLICATION NUMBER: US/09/633,364
;  CURRENT FILING DATE:  2000-08-07
;  PRIOR APPLICATION NUMBER: US 60/171,312
;  PRIOR FILING DATE: 1999-12-21
;  NUMBER OF SEQ ID NOS: 6899
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5244
;   LENGTH: 8
;   TYPE: PRT
;   ORGANISM: Homo sapien
US-09-633-364-5244


  Query Match             41.1%;  Score 23;  DB 5;  Length 8;
  Best Local Similarity   80.0%;  Pred. No. 2.9e+05;
  Matches    4;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 CMTWN 5
          ||| |
Db      2 CMTTN 6


RESULT  15
US-09-922-227-8
; Sequence 8, Application US/09922227
;   GENERAL INFORMATION:
;        APPLICANT: Ruoslahti, Erkki
;                   Pasqualini, Renata
;        TITLE OF INVENTION: Method of Identifying Molecules That
;                            Home to a Selected Organ In Vivo
;        NUMBER OF SEQUENCES: 44
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Campbell & Flores LLP
;             STREET: 4370 La Jolla Village Drive, Suite 700
;             CITY: San Diego
;             STATE: California
;             COUNTRY: United States
;             ZIP: 92122
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.25
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/922,227
;             FILING DATE: 02-Aug-2001
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US 08/526,710
;             FILING DATE: 11-SEP-1995
;             APPLICATION NUMBER: US 08/813,273
;             FILING DATE: 10-MAR-1997
;             APPLICATION NUMBER: US 08/862,855
;             FILING DATE: 23-MAY-1997
;             APPLICATION NUMBER: US 09/227,906
;             FILING DATE: 08-JAN-1999
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Campbell, Cathryn A.
;             REGISTRATION NUMBER: 31,815
;             REFERENCE/DOCKET NUMBER: P-LJ 4859
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (619) 535-9001
;             TELEFAX: (619) 535-8949
;   INFORMATION FOR SEQ ID NO: 8:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 8 amino acids
;             TYPE: amino acid
;             TOPOLOGY: linear
;        MOLECULE TYPE: peptide
;        SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-09-922-227-8

  Query Match             41.1%;  Score 23;  DB 5;  Length 8;
  Best Local Similarity   28.6%;  Pred. No. 2.9e+05;
  Matches    2;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 CMTWNQM 7
          |: | ::
Db      1 CLDWGRI 7


Search completed: October 23, 2001, 13:30:23
Job time: 145 sec
```

```
;  APPLICANT:  Sleath, Paul R.
;  APPLICANT:  Mossman, Sally
;  APPLICANT:  Evans, Lawrence
;  APPLICANT:  Spies, A. Gregory
;  APPLICANT:  Boydston, Jeremy
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1
;  TITLE OF INVENTION:  SPECIFIC IMMUNOTHERAPY
;  FILE REFERENCE: 210121.465C5
;  CURRENT APPLICATION NUMBER: US/09/938,864
;  CURRENT FILING DATE:  2001-08-24
;  NUMBER OF SEQ ID NOS: 413
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 286
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-938-864-286


  Query Match             46.4%;  Score 26;  DB 5;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.9e+05;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          5 NQMNL 9
              |||||
Db          1 NQMNL 5


RESULT  11
US-09-832-723-108
; Sequence 108, Application US/09832723
; GENERAL INFORMATION:
;  APPLICANT: Estell, David A.
;  APPLICANT:  Chen, Yiyou
;  APPLICANT:  Murray, Christopher J.
;  APPLICANT:  Tijerina, Pilar
;  TITLE OF INVENTION: METHODS FOR SELECTIVE TARGETING
;  FILE REFERENCE: GC617-2
;  CURRENT APPLICATION NUMBER: US/09/832,723
;  CURRENT FILING DATE:  2001-04-11
;  PRIOR APPLICATION NUMBER: US 60/197,259
;  PRIOR FILING DATE: 2000-04-14
;  NUMBER OF SEQ ID NOS: 117
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 108
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: peptides screened from a phage display random
;   OTHER INFORMATION: peptide library
US-09-832-723-108


  Query Match             42.9%;  Score 24;  DB 5;  Length 9;
  Best Local Similarity   60.0%;  Pred. No. 2.9e+05;
  Matches    3;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 CMTWN 5
              |  ||
Db          1 CAPWN 5


RESULT  12
US-09-832-723-109
; Sequence 109, Application US/09832723
; GENERAL INFORMATION:
;  APPLICANT: Estell, David A.
;  APPLICANT:  Chen, Yiyou
;  APPLICANT:  Murray, Christopher J.
;  APPLICANT:  Tijerina, Pilar
;  TITLE OF INVENTION: METHODS FOR SELECTIVE TARGETING
;  FILE REFERENCE: GC617-2
```

```
;  CURRENT APPLICATION NUMBER: US/09/832,723
;  CURRENT FILING DATE:  2001-04-11
;  PRIOR APPLICATION NUMBER: US 60/197,259
;  PRIOR FILING DATE: 2000-04-14
;  NUMBER OF SEQ ID NOS: 117
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 109
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: peptides screened from a phage display random
;   OTHER INFORMATION: peptide library
US-09-832-723-109

  Query Match             42.9%;  Score 24;  DB 5;  Length 9;
  Best Local Similarity   60.0%;  Pred. No. 2.9e+05;
  Matches    3;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 CMTWN 5
              |  ||
Db          1 CAPWN 5


RESULT  13
US-09-633-364-4214
; Sequence 4214, Application US/09633364
; GENERAL INFORMATION:
;  APPLICANT: Fikes, John
;  APPLICANT:  Sette, Alessandro
;  APPLICANT:  Sidney, John
;  APPLICANT:  Southwood, Scott
;  APPLICANT:  Chesnut, Robert
;  APPLICANT:  Celis, Esteban
;  APPLICANT:  Keogh, Elissa
;  APPLICANT:  Epimmune Inc.
;  TITLE OF INVENTION: Inducing Cellular Immune Responses to
;  TITLE OF INVENTION:  Prostate Cancer Antigens Using Peptide and Nucleic Acid
;  TITLE OF INVENTION:  Compositions
;  FILE REFERENCE: 018623-014710US
;  CURRENT APPLICATION NUMBER: US/09/633,364
;  CURRENT FILING DATE:  2000-08-07
;  PRIOR APPLICATION NUMBER: US 60/171,312
;  PRIOR FILING DATE: 1999-12-21
;  NUMBER OF SEQ ID NOS: 6899
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4214
;   LENGTH: 8
;   TYPE: PRT
;   ORGANISM: Homo sapien
US-09-633-364-4214


  Query Match             41.1%;  Score 23;  DB 5;  Length 8;
  Best Local Similarity   80.0%;  Pred. No. 2.9e+05;
  Matches    4;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 CMTWN 5
              ||| |
Db          2 CMTTN 6


RESULT  14
US-09-633-364-5244
; Sequence 5244, Application US/09633364
; GENERAL INFORMATION:
;  APPLICANT: Fikes, John
;  APPLICANT:  Sette, Alessandro
;  APPLICANT:  Sidney, John
;  APPLICANT:  Southwood, Scott
;  APPLICANT:  Chesnut, Robert
;  APPLICANT:  Celis, Esteban
```

```
;  APPLICANT:  Evans, Lawrence
;  APPLICANT:  Spies, A. Gregory
;  APPLICANT:  Boydston, Jeremy
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1
;  TITLE OF INVENTION:  SPECIFIC IMMUNOTHERAPY
;  FILE REFERENCE: 210121.465C5
;  CURRENT APPLICATION NUMBER: US/09/938,864
;  CURRENT FILING DATE:  2001-08-24
;  NUMBER OF SEQ ID NOS: 413
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 245
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Homo sapien
US-09-938-864-245

  Query Match            66.1%;  Score 37;  DB 5;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.9e+05;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       4 WNQMNL 9
           ||||||
Db       1 WNQMNL 6

RESULT   7
US-09-938-864-305
; Sequence 305, Application US/09938864
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Smithgall, Molly
;  APPLICANT:  Moulton, Gus
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Sleath, Paul R.
;  APPLICANT:  Mossman, Sally
;  APPLICANT:  Evans, Lawrence
;  APPLICANT:  Spies, A. Gregory
;  APPLICANT:  Boydston, Jeremy
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1
;  TITLE OF INVENTION:  SPECIFIC IMMUNOTHERAPY
;  FILE REFERENCE: 210121.465C5
;  CURRENT APPLICATION NUMBER: US/09/938,864
;  CURRENT FILING DATE:  2001-08-24
;  NUMBER OF SEQ ID NOS: 413
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 305
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-938-864-305

  Query Match            66.1%;  Score 37;  DB 5;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.9e+05;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       4 WNQMNL 9
           ||||||
Db       1 WNQMNL 6

RESULT   8
US-09-938-864-234
; Sequence 234, Application US/09938864
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Smithgall, Molly
;  APPLICANT:  Moulton, Gus
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Sleath, Paul R.
```

```
;  APPLICANT:  Mossman, Sally
;  APPLICANT:  Evans, Lawrence
;  APPLICANT:  Spies, A. Gregory
;  APPLICANT:  Boydston, Jeremy
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1
;  TITLE OF INVENTION:  SPECIFIC IMMUNOTHERAPY
;  FILE REFERENCE: 210121.465C5
;  CURRENT APPLICATION NUMBER: US/09/938,864
;  CURRENT FILING DATE:  2001-08-24
;  NUMBER OF SEQ ID NOS: 413
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 234
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Homo sapien
US-09-938-864-234

  Query Match            53.6%;  Score 30;  DB 5;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.9e+05;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CMTW 4
           ||||
Db       6 CMTW 9

RESULT   9
US-09-938-864-151
; Sequence 151, Application US/09938864
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Smithgall, Molly
;  APPLICANT:  Moulton, Gus
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Sleath, Paul R.
;  APPLICANT:  Mossman, Sally
;  APPLICANT:  Evans, Lawrence
;  APPLICANT:  Spies, A. Gregory
;  APPLICANT:  Boydston, Jeremy
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1
;  TITLE OF INVENTION:  SPECIFIC IMMUNOTHERAPY
;  FILE REFERENCE: 210121.465C5
;  CURRENT APPLICATION NUMBER: US/09/938,864
;  CURRENT FILING DATE:  2001-08-24
;  NUMBER OF SEQ ID NOS: 413
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 151
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Homo sapien
US-09-938-864-151

  Query Match            46.4%;  Score 26;  DB 5;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.9e+05;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       5 NQMNL 9
           |||||
Db       1 NQMNL 5

RESULT  10
US-09-938-864-286
; Sequence 286, Application US/09938864
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Smithgall, Molly
;  APPLICANT:  Moulton, Gus
;  APPLICANT:  Vedvick, Thomas S.
```

```
;  APPLICANT:  Boydston, Jeremy
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1
;  TITLE OF INVENTION:  SPECIFIC IMMUNOTHERAPY
;  FILE REFERENCE: 210121.465C5
;  CURRENT APPLICATION NUMBER: US/09/938,864
;  CURRENT FILING DATE:  2001-08-24
;  NUMBER OF SEQ ID NOS: 413
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 258
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-938-864-258

  Query Match             100.0%;  Score 56;  DB 5;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.9e+05;
  Matches '  9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CMTWNQMNL 9
           |||||||||
Db       1 CMTWNQMNL 9

RESULT   3
US-09-938-864-131
; Sequence 131, Application US/09938864
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Smithgall, Molly
;  APPLICANT:  Moulton, Gus
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Sleath, Paul R.
;  APPLICANT:  Mossman, Sally
;  APPLICANT:  Evans, Lawrence
;  APPLICANT:  Spies, A. Gregory
;  APPLICANT:  Boydston, Jeremy
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1
;  TITLE OF INVENTION:  SPECIFIC IMMUNOTHERAPY
;  FILE REFERENCE: 210121.465C5
;  CURRENT APPLICATION NUMBER: US/09/938,864
;  CURRENT FILING DATE:  2001-08-24
;  NUMBER OF SEQ ID NOS: 413
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 131
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Homo sapien
US-09-938-864-131

  Query Match             82.1%;  Score 46;  DB 5;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.9e+05;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CMTWNQM 7
           |||||||
Db       3 CMTWNQM 9

RESULT   4
US-09-938-864-278
; Sequence 278, Application US/09938864
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Smithgall, Molly
;  APPLICANT:  Moulton, Gus
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Sleath, Paul R.
;  APPLICANT:  Mossman, Sally
;  APPLICANT:  Evans, Lawrence
;  APPLICANT:  Spies, A. Gregory
;  APPLICANT:  Boydston, Jeremy
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1
;  TITLE OF INVENTION:  SPECIFIC IMMUNOTHERAPY
;  FILE REFERENCE: 210121.465C5
;  CURRENT APPLICATION NUMBER: US/09/938,864
;  CURRENT FILING DATE:  2001-08-24
;  NUMBER OF SEQ ID NOS: 413
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 278
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-938-864-278

  Query Match             82.1%;  Score 46;  DB 5;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.9e+05;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CMTWNQM 7
           |||||||
Db       3 CMTWNQM 9

RESULT   5
US-09-938-864-167
; Sequence 167, Application US/09938864
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Smithgall, Molly
;  APPLICANT:  Moulton, Gus
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Sleath, Paul R.
;  APPLICANT:  Mossman, Sally
;  APPLICANT:  Evans, Lawrence
;  APPLICANT:  Spies, A. Gregory
;  APPLICANT:  Boydston, Jeremy
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1
;  TITLE OF INVENTION:  SPECIFIC IMMUNOTHERAPY
;  FILE REFERENCE: 210121.465C5
;  CURRENT APPLICATION NUMBER: US/09/938,864
;  CURRENT FILING DATE:  2001-08-24
;  NUMBER OF SEQ ID NOS: 413
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 167
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Homo sapien
US-09-938-864-167

  Query Match             73.2%;  Score 41;  DB 5;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.9e+05;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CMTWNQ 6
           ||||||
Db       4 CMTWNQ 9

RESULT   6
US-09-938-864-245
; Sequence 245, Application US/09938864
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Smithgall, Molly
;  APPLICANT:  Moulton, Gus
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Sleath, Paul R.
;  APPLICANT:  Mossman, Sally
```

```
                       GenCore version 4.5
              Copyright (c) 1993 - 2000  Compugen Ltd.

OM protein - protein search, using sw model

Run on:          October 23, 2001, 13:30:23 ; Search time 26.91 Seconds
                                              (without alignments)
                                              17.233 Million cell updates/sec

Title:           US-09-625-963-2
Perfect score:   56
Sequence:        1 CMTWNQMNL 9

Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched:        339663 seqs, 51527020 residues

Total number of hits satisfying chosen parameters:      53587

Minimum DB seq length: 0
Maximum DB seq length: 9

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       Pending_Patents_AA_New:*
                1:  /cgn2_6/ptodata/2/paa/PCT_NEW_COMB.pep:*
                2:  /cgn2_6/ptodata/2/paa/US06_NEW_COMB.pep:*
                3:  /cgn2_6/ptodata/2/paa/US07_NEW_COMB.pep:*
                4:  /cgn2_6/ptodata/2/paa/US08_NEW_COMB.pep:*
                5:  /cgn2_6/ptodata/2/paa/US09_NEW_COMB.pep:*
                6:  /cgn2_6/ptodata/2/paa/US60_NEW_COMB.pep:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                                  SUMMARIES
                      %
Result              Query
   No.     Score  Match Length DB  ID                          Description
 ------------------------------------------------------------------------------
     1       56   100.0      9  5  US-09-938-864-49            Sequence 49, Appl
     2       56   100.0      9  5  US-09-938-864-258           Sequence 258, App
     3       46    82.1      9  5  US-09-938-864-131           Sequence 131, App
     4       46    82.1      9  5  US-09-938-864-278           Sequence 278, App
     5       41    73.2      9  5  US-09-938-864-167           Sequence 167, App
     6       37    66.1      9  5  US-09-938-864-245           Sequence 245, App
     7       37    66.1      9  5  US-09-938-864-305           Sequence 305, App
     8       30    53.6      9  5  US-09-938-864-234           Sequence 234, App
     9       26    46.4      9  5  US-09-938-864-151           Sequence 151, App
    10       26    46.4      9  5  US-09-938-864-286           Sequence 286, App
    11       24    42.9      9  5  US-09-832-723-108           Sequence 108, App
    12       24    42.9      9  5  US-09-832-723-109           Sequence 109, App
    13       23    41.1      8  5  US-09-633-364-4214          Sequence 4214, Ap
    14       23    41.1      8  5  US-09-633-364-5244          Sequence 5244, Ap
    15       23    41.1      8  5  US-09-922-227-8             Sequence 8, Appli
    16       23    41.1      9  5  US-09-631-863A-87           Sequence 87, Appl
    17       23    41.1      9  5  US-09-633-364-6478          Sequence 6478, Ap
    18       23    41.1      9  5  US-09-780-053-131           Sequence 131, App
    19       23    41.1      9  5  US-09-780-053-252           Sequence 252, App
    20       23    41.1      9  5  US-09-780-053-457           Sequence 457, App
    21       23    41.1      9  5  US-09-780-053-547           Sequence 547, App
    22       23    41.1      9  5  US-09-809-638-121           Sequence 121, App
    23       23    41.1      9  5  US-09-809-638-262           Sequence 262, App
    24       22    39.3      8  5  US-09-350-641B-1621         Sequence 1621, Ap
    25       22    39.3      9  5  US-09-533-798-20            Sequence 20, Appl
    26       22    39.3      9  5  US-09-793-451-454           Sequence 454, App
    27       21    37.5      7  5  US-09-922-227-29            Sequence 29, Appl
    28       21    37.5      8  5  US-09-548-936B-5            Sequence 5, Appli
    29       21    37.5      9  5  US-09-786-214-42            Sequence 42, Appl
    30       21    37.5      9  5  US-09-766-889A-42           Sequence 42, Appl
    31       21    37.5      9  5  US-09-839-542-2741          Sequence 2741, Ap
    32       21    37.5      9  5  US-09-806-769-26            Sequence 26, Appl
    33       21    37.5      9  5  US-09-688-990-22            Sequence 22, Appl
    34       21    37.5      9  5  US-09-923-831-30            Sequence 30, Appl
    35       20    35.7      6  5  US-09-620-978-569           Sequence 569, App
    36       20    35.7      6  5  US-09-570-581A-392          Sequence 392, App
    37       20    35.7      6  5  US-09-602-025A-392          Sequence 392, App
    38       20    35.7      6  5  US-09-606-181A-392          Sequence 392, App
    39       20    35.7      6  5  US-09-596-577A-392          Sequence 392, App
    40       20    35.7      6  5  US-09-620-421-392           Sequence 392, App
    41       20    35.7      6  5  US-09-635-277A-569          Sequence 569, App
    42       20    35.7      6  5  US-09-637-563A-569          Sequence 569, App
    43       20    35.7      6  5  US-09-640-695A-569          Sequence 569, App
    44       20    35.7      6  5  US-09-667-597A-569          Sequence 569, App
    45       20    35.7      6  5  US-09-570-768A-408          Sequence 408, App

                                  ALIGNMENTS


RESULT   1
US-09-938-864-49
; Sequence 49, Application US/09938864
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Smithgall, Molly
;  APPLICANT:  Moulton, Gus
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Sleath, Paul R.
;  APPLICANT:  Mossman, Sally
;  APPLICANT:  Evans, Lawrence
;  APPLICANT:  Spies, A. Gregory
;  APPLICANT:  Boydston, Jeremy
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1
;  TITLE OF INVENTION:  SPECIFIC IMMUNOTHERAPY
;  FILE REFERENCE: 210121.465C5
;  CURRENT APPLICATION NUMBER: US/09/938,864
;  CURRENT FILING DATE:  2001-08-24
;  NUMBER OF SEQ ID NOS: 413
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 49
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Homo sapien
US-09-938-864-49


  Query Match             100.0%;  Score 56;  DB 5;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.9e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 CMTWNQMNL 9
            |||||||||
Db        1 CMTWNQMNL 9


RESULT   2
US-09-938-864-258
; Sequence 258, Application US/09938864
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Smithgall, Molly
;  APPLICANT:  Moulton, Gus
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Sleath, Paul R.
;  APPLICANT:  Mossman, Sally
;  APPLICANT:  Evans, Lawrence
;  APPLICANT:  Spies, A. Gregory
```

```
;  TITLE OF INVENTION: METHODS FOR THE DIAGNOSIS AND THERAPY
;  TITLE OF INVENTION: OF CHLAMYDIA INFECTION
;  FILE REFERENCE: 210121.492
;  CURRENT APPLICATION NUMBER: US/09/523,419
;  CURRENT FILING DATE: 2000-03-10
;  NUMBER OF SEQ ID NOS: 326
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 258
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-523-419-258


  Query Match             100.0%;  Score 56;  DB 19;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.6e+06;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 CMTWNQMNL 9
            |||||||||
Db        1 CMTWNQMNL 9

RESULT  13
US-09-625-963-2
; Sequence 2, Application US/09625963
; GENERAL INFORMATION:
;  APPLICANT:  Strauss, Hans Josef
;  APPLICANT:  Gao, Liquan
;  TITLE OF INVENTION:  Immunotherapeutic Methods Using Epitopes of WT-1 and
;  TITLE OF INVENTION:  GATA-1
;  FILE REFERENCE:  ICI 101
;  CURRENT APPLICATION NUMBER: US/09/625,963
;  CURRENT FILING DATE:  2000-07-26
;  PRIOR APPLICATION NUMBER:  PCT/GB99/03572
;  PRIOR FILING DATE:  1999-11-02
;  PRIOR APPLICATION NUMBER:  GB9823897.5
;  PRIOR FILING DATE:  1998-11-02
;  NUMBER OF SEQ ID NOS:  5
;  SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 2
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-625-963-2


  Query Match             100.0%;  Score 56;  DB 20;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.6e+06;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 CMTWNQMNL 9
            |||||||||
Db        1 CMTWNQMNL 9

RESULT  14
US-09-679-339-49
; Sequence 49, Application US/09679339
; GENERAL INFORMATION:
;  APPLICANT:  Skeiky, Yasir
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Cheever, Martin
;  APPLICANT:  Reed, Steven G.
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR WT1
;  TITLE OF INVENTION:  SPECIFIC IMMUNOTHERAPY
;  FILE REFERENCE:  210121.533
;  CURRENT APPLICATION NUMBER: US/09/679,339
;  CURRENT FILING DATE:  2000-10-04
;  NUMBER OF SEQ ID NOS:  357
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 49
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Homo sapien
US-09-679-339-49


  Query Match             100.0%;  Score 56;  DB 20;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.6e+06;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 CMTWNQMNL 9
            |||||||||
Db        1 CMTWNQMNL 9

RESULT  15
US-09-679-339-258
; Sequence 258, Application US/09679339
; GENERAL INFORMATION:
;  APPLICANT:  Skeiky, Yasir
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Cheever, Martin
;  APPLICANT:  Reed, Steven G.
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR WT1
;  TITLE OF INVENTION:  SPECIFIC IMMUNOTHERAPY
;  FILE REFERENCE:  210121.533
;  CURRENT APPLICATION NUMBER: US/09/679,339
;  CURRENT FILING DATE:  2000-10-04
;  NUMBER OF SEQ ID NOS:  357
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 258
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-679-339-258


  Query Match             100.0%;  Score 56;  DB 20;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.6e+06;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 CMTWNQMNL 9
            |||||||||
Db        1 CMTWNQMNL 9


Search completed: October 23, 2001, 13:33:06
Job time: 283 sec
```

```
;  NUMBER OF SEQ ID NOS: 326
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 49
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Homo sapien
US-09-276-484-49

  Query Match              100.0%;  Score 56;  DB 16;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.6e+06;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 CMTWNQMNL 9
            |||||||||
Db        1 CMTWNQMNL 9

RESULT   8
US-09-276-484-258
; Sequence 258, Application US/09276484
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
;  APPLICANT: Cheever, Martin A.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1
;  TITLE OF INVENTION: SPECIFIC IMMUNOTHERAPY
;  FILE REFERENCE: 210121.465C1
;  CURRENT APPLICATION NUMBER: US/09/276,484
;  CURRENT FILING DATE: 1999-03-25
;  NUMBER OF SEQ ID NOS: 326
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 258
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-276-484-258

  Query Match              100.0%;  Score 56;  DB 16;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.6e+06;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 CMTWNQMNL 9
            |||||||||
Db        1 CMTWNQMNL 9

RESULT   9
US-09-276-484A-49
; Sequence 49, Application US/09276484A
; GENERAL INFORMATION:
;  APPLICANT:  Gaiger, Alexander
;  APPLICANT:  Cheever, Martin A.
;  TITLE OF INVENTION:  COMPOSITONS AND METHODS FOR WT1 SPECIFIC
;  TITLE OF INVENTION:  IMMUNOTHERAPY
;  FILE REFERENCE:  210121.465C1
;  CURRENT APPLICATION NUMBER: US/09/276,484A
;  CURRENT FILING DATE:  1999-03-25
;  PRIOR APPLICATION NUMBER:  US 09/164,223
;  PRIOR FILING DATE:  1999-09-30
;  NUMBER OF SEQ ID NOS:  326
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 49
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Homo sapien
US-09-276-484A-49

  Query Match              100.0%;  Score 56;  DB 16;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.6e+06;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 CMTWNQMNL 9
            |||||||||
Db        1 CMTWNQMNL 9

RESULT  10
US-09-276-484A-258
; Sequence 258, Application US/09276484A
; GENERAL INFORMATION:
;  APPLICANT:  Gaiger, Alexander
;  APPLICANT:  Cheever, Martin A.
;  TITLE OF INVENTION:  COMPOSITONS AND METHODS FOR WT1 SPECIFIC
;  TITLE OF INVENTION:  IMMUNOTHERAPY
;  FILE REFERENCE:  210121.465C1
;  CURRENT APPLICATION NUMBER: US/09/276,484A
;  CURRENT FILING DATE:  1999-03-25
;  PRIOR APPLICATION NUMBER:  US 09/164,223
;  PRIOR FILING DATE:  1999-09-30
;  NUMBER OF SEQ ID NOS:  326
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 258
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-276-484A-258

  Query Match              100.0%;  Score 56;  DB 16;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.6e+06;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 CMTWNQMNL 9
            |||||||||
Db        1 CMTWNQMNL 9

RESULT  11
US-09-523-419-49
; Sequence 49, Application US/09523419
; GENERAL INFORMATION:
;  APPLICANT: Hendrickson, Ronald C.
;  APPLICANT: Probst, Peter
;  APPLICANT: Reed, Steven G.
;  TITLE OF INVENTION: METHODS FOR THE DIAGNOSIS AND THERAPY
;  TITLE OF INVENTION: OF CHLAMYDIA INFECTION
;  FILE REFERENCE: 210121.492
;  CURRENT APPLICATION NUMBER: US/09/523,419
;  CURRENT FILING DATE: 2000-03-10
;  NUMBER OF SEQ ID NOS: 326
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 49
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Homo sapien
US-09-523-419-49

  Query Match              100.0%;  Score 56;  DB 19;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.6e+06;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 CMTWNQMNL 9
            |||||||||
Db        1 CMTWNQMNL 9

RESULT  12
US-09-523-419-258
; Sequence 258, Application US/09523419
; GENERAL INFORMATION:
;  APPLICANT: Hendrickson, Ronald C.
;  APPLICANT: Probst, Peter
;  APPLICANT: Reed, Steven G.
```

```
;  APPLICANT: Cheever, Martin A.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1 SPECIFIC IMMUNOTHERAPY
;  FILE REFERENCE: 210121.465
;  CURRENT APPLICATION NUMBER: US/09/164,223
;  CURRENT FILING DATE: 1998-09-30
;  NUMBER OF SEQ ID NOS: 320
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 258
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-164-223-258


  Query Match             100.0%;  Score 56;  DB 15;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.6e+06;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CMTWNQMNL 9
           |||||||||
Db       1 CMTWNQMNL 9

RESULT   3
US-09-164-223-49
; Sequence 49, Application US/09164223A
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
;  APPLICANT: Cheever, Martin A.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1 SPECIFIC IMMUNOTHERAPY
;  FILE REFERENCE: 210121.465
;  CURRENT APPLICATION NUMBER: US/09/164,223A
;  CURRENT FILING DATE: 1998-09-30
;  NUMBER OF SEQ ID NOS: 320
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 49
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-164-223-49


  Query Match             100.0%;  Score 56;  DB 15;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.6e+06;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CMTWNQMNL 9
           |||||||||
Db       1 CMTWNQMNL 9

RESULT   4
US-09-164-223-258
; Sequence 258, Application US/09164223A
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
;  APPLICANT: Cheever, Martin A.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1 SPECIFIC IMMUNOTHERAPY
;  FILE REFERENCE: 210121.465
;  CURRENT APPLICATION NUMBER: US/09/164,223A
;  CURRENT FILING DATE: 1998-09-30
;  NUMBER OF SEQ ID NOS: 320
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 258
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-164-223-258


  Query Match             100.0%;  Score 56;  DB 15;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.6e+06;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy       1 CMTWNQMNL 9
           |||||||||
Db       1 CMTWNQMNL 9

RESULT   5
US-09-164-223-49
; Sequence 49, Application US/09164223B
; GENERAL INFORMATION:
;  APPLICANT:  Gaiger, Alexander
;  APPLICANT:  Cheever, Martin A.
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR WT1 SPECIFIC IMMUNOTHERAPY
;  FILE REFERENCE:  210121.465
;  CURRENT APPLICATION NUMBER: US/09/164,223B
;  CURRENT FILING DATE:  1998-09-30
;  NUMBER OF SEQ ID NOS:  320
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 49
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-164-223-49


  Query Match             100.0%;  Score 56;  DB 15;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.6e+06;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CMTWNQMNL 9
           |||||||||
Db       1 CMTWNQMNL 9

RESULT   6
US-09-164-223-258
; Sequence 258, Application US/09164223B
; GENERAL INFORMATION:
;  APPLICANT:  Gaiger, Alexander
;  APPLICANT:  Cheever, Martin A.
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR WT1 SPECIFIC IMMUNOTHERAPY
;  FILE REFERENCE:  210121.465
;  CURRENT APPLICATION NUMBER: US/09/164,223B
;  CURRENT FILING DATE:  1998-09-30
;  NUMBER OF SEQ ID NOS:  320
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 258
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-164-223-258


  Query Match             100.0%;  Score 56;  DB 15;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.6e+06;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CMTWNQMNL 9
           |||||||||
Db       1 CMTWNQMNL 9

RESULT   7
US-09-276-484-49
; Sequence 49, Application US/09276484
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
;  APPLICANT: Cheever, Martin A.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1
;  TITLE OF INVENTION: SPECIFIC IMMUNOTHERAPY
;  FILE REFERENCE: 210121.465C1
;  CURRENT APPLICATION NUMBER: US/09/276,484
;  CURRENT FILING DATE: 1999-03-25
```

```
                        GenCore version 4.5
              Copyright (c) 1993 - 2000  Compugen Ltd.

OM protein - protein search, using sw model

Run on:         October 23, 2001, 13:33:06 ; Search time 155.83 Seconds
                                             (without alignments)
                                             12.813 Million cell updates/sec

Title:          US-09-625-963-2
Perfect score:  56
Sequence:       1 CMTWNQMNL 9

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched:       2803329 seqs, 221847457 residues

Total number of hits satisfying chosen parameters:      162356

Minimum DB seq length: 0
Maximum DB seq length: 9

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :      Pending_Patents_AA_Main:*
               1:  /cgn2_6/ptodata/2/paa/PCTUS_COMB.pep:*
               2:  /cgn2_6/ptodata/2/paa/US06_COMB.pep:*
               3:  /cgn2_6/ptodata/2/paa/US07_COMB.pep:*
               4:  /cgn2_6/ptodata/2/paa/US080_COMB.pep:*
               5:  /cgn2_6/ptodata/2/paa/US081_COMB.pep:*
               6:  /cgn2_6/ptodata/2/paa/US082_COMB.pep:*
               7:  /cgn2_6/ptodata/2/paa/US083_COMB.pep:*
               8:  /cgn2_6/ptodata/2/paa/US084_COMB.pep:*
               9:  /cgn2_6/ptodata/2/paa/US085_COMB.pep:*
               10:  /cgn2_6/ptodata/2/paa/US086_COMB.pep:*
               11:  /cgn2_6/ptodata/2/paa/US087_COMB.pep:*
               12:  /cgn2_6/ptodata/2/paa/US088_COMB.pep:*
               13:  /cgn2_6/ptodata/2/paa/US089_COMB.pep:*
               14:  /cgn2_6/ptodata/2/paa/US090_COMB.pep:*
               15:  /cgn2_6/ptodata/2/paa/US091_COMB.pep:*
               16:  /cgn2_6/ptodata/2/paa/US092_COMB.pep:*
               17:  /cgn2_6/ptodata/2/paa/US093_COMB.pep:*
               18:  /cgn2_6/ptodata/2/paa/US094_COMB.pep:*
               19:  /cgn2_6/ptodata/2/paa/US095_COMB.pep:*
               20:  /cgn2_6/ptodata/2/paa/US096_COMB.pep:*
               21:  /cgn2_6/ptodata/2/paa/US097_COMB.pep:*
               22:  /cgn2_6/ptodata/2/paa/US098_COMB.pep:*
               23:  /cgn2_6/ptodata/2/paa/US60_COMB.pep:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                                 SUMMARIES
                   %
Result           Query
  No.   Score    Match Length DB  ID                    Description
-----------------------------------------------------------------------
    1      56    100.0      9  15  US-09-164-223-49          Sequence 49, Appl
    2      56    100.0      9  15  US-09-164-223-258         Sequence 258, App
    3      56    100.0      9  15  US-09-164-223-49          Sequence 49, Appl
    4      56    100.0      9  15  US-09-164-223-258         Sequence 258, App
    5      56    100.0      9  15  US-09-164-223-49          Sequence 49, Appl
    6      56    100.0      9  15  US-09-164-223-258         Sequence 258, App
    7      56    100.0      9  16  US-09-276-484-49          Sequence 49, Appl
    8      56    100.0      9  16  US-09-276-484-258         Sequence 258, App
    9      56    100.0      9  16  US-09-276-484A-49         Sequence 49, Appl
   10      56    100.0      9  16  US-09-276-484A-258        Sequence 258, App
   11      56    100.0      9  19  US-09-523-419-49          Sequence 49, Appl
   12      56    100.0      9  19  US-09-523-419-258         Sequence 258, App
   13      56    100.0      9  20  US-09-625-963-2           Sequence 2, Appli
   14      56    100.0      9  20  US-09-679-339-49          Sequence 49, Appl
   15      56    100.0      9  20  US-09-679-339-258         Sequence 258, App
   16      56    100.0      9  20  US-09-684-361-49          Sequence 49, Appl
   17      56    100.0      9  20  US-09-684-361-258         Sequence 258, App
   18      56    100.0      9  20  US-09-685-830-49          Sequence 49, Appl
   19      56    100.0      9  20  US-09-685-830-258         Sequence 258, App
   20      56    100.0      9  21  US-09-785-019-49          Sequence 49, Appl
   21      56    100.0      9  21  US-09-785-019-258         Sequence 258, App
   22      56    100.0      9  21  US-09-791-477-49          Sequence 49, Appl
   23      56    100.0      9  21  US-09-791-477-258         Sequence 258, App
   24      46     82.1      9  15  US-09-164-223-131         Sequence 131, App
   25      46     82.1      9  15  US-09-164-223-278         Sequence 278, App
   26      46     82.1      9  15  US-09-164-223-131         Sequence 131, App
   27      46     82.1      9  15  US-09-164-223-278         Sequence 278, App
   28      46     82.1      9  15  US-09-164-223-131         Sequence 131, App
   29      46     82.1      9  15  US-09-164-223-278         Sequence 278, App
   30      46     82.1      9  16  US-09-276-484-131         Sequence 131, App
   31      46     82.1      9  16  US-09-276-484-278         Sequence 278, App
   32      46     82.1      9  16  US-09-276-484A-131        Sequence 131, App
   33      46     82.1      9  16  US-09-276-484A-278        Sequence 278, App
   34      46     82.1      9  19  US-09-523-419-131         Sequence 131, App
   35      46     82.1      9  19  US-09-523-419-278         Sequence 278, App
   36      46     82.1      9  20  US-09-679-339-131         Sequence 131, App
   37      46     82.1      9  20  US-09-679-339-278         Sequence 278, App
   38      46     82.1      9  20  US-09-684-361-131         Sequence 131, App
   39      46     82.1      9  20  US-09-684-361-278         Sequence 278, App
   40      46     82.1      9  20  US-09-685-830-131         Sequence 131, App
   41      46     82.1      9  20  US-09-685-830-278         Sequence 278, App
   42      46     82.1      9  21  US-09-785-019-131         Sequence 131, App
   43      46     82.1      9  21  US-09-785-019-278         Sequence 278, App
   44      46     82.1      9  21  US-09-791-477-131         Sequence 131, App
   45      46     82.1      9  21  US-09-791-477-278         Sequence 278, App

                                ALIGNMENTS


RESULT   1
US-09-164-223-49
; Sequence 49, Application US/09164223
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
;  APPLICANT: Cheever, Martin A.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1 SPECIFIC IMMUNOTHERAPY
;  FILE REFERENCE: 210121.465
;  CURRENT APPLICATION NUMBER: US/09/164,223
;  CURRENT FILING DATE: 1998-09-30
;  NUMBER OF SEQ ID NOS: 320
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 49
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-164-223-49


  Query Match             100.0%;  Score 56;  DB 15;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.6e+06;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 CMTWNQMNL 9
            |||||||||
Db        1 CMTWNQMNL 9


RESULT   2
US-09-164-223-258
; Sequence 258, Application US/09164223
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
```

```
FT                      /note= "D-form residue"
FT   Misc-difference 3
FT                      /note= "D-form residue"
FT   Misc-difference 4
FT                      /note= "D-form residue"
FT   Misc-difference 6
FT                      /note= "D-form residue, in C-terminal amide form"
XX
PN   WO9640208-A1.
XX
PD   19-DEC-1996.
XX
PF   06-JUN-1996;   96WO-US09321.
XX
PR   07-JUN-1995;   95US-0476438.
XX
PA   (TORR-) TORREY PINES INST MOLECULAR STUDIES.
XX
PI   Dooley CT,  Houghten RA;
XX
DR   WPI; 1997-051895/05.
XX
PT   New mu opioid receptor binding ligand peptide(s) - useful for
PT   in-vitro and in-vivo diagnosis, as analgesics, and for blocking
PT   peripheral effects of centrally acting drugs, e.g. morphine
XX
PS   Disclosure; Page 26; 57pp; English.
XX
CC   The patent discloses eight new groups of opioid peptides which bind
CC   to the mu-receptor to act as agonists or antagonists. The peptides
CC   can be used for in-vitro assays to study opiate receptor subtypes
CC   (especially the mu type) in brain or other tissue samples; and for
CC   in-vivo diagnosis to localise opioid subtypes. The peptides are also
CC   useful as drugs to treat pathologies associated with other compounds
CC   which interact with the opioid receptor system. Therefore they can be
CC   used in medicaments for treating pathologies associated with the mu
CC   receptor and as analgesics. They can be used therapeutically to block
CC   the peripheral effects of centrally acting pain killers, e.g. to
CC   prevent side effects such as constipation and pruritis associated
CC   with morphine. The present sequence represents a specific example
CC   of one of the new groups of peptides, of formula
CC   (D)Ile-(D)Met-(D)Thr-(D)Trp-Gly-Xaa-NH2 where Xaa = Gly or the D-form
CC   of a naturally occurring amino acid.
XX
SQ   Sequence   6 AA;


  Query Match             46.4%;  Score 26;  DB 18;  Length 6;
  Best Local Similarity   80.0%;  Pred. No. 3.4e+05;
  Matches    4;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 MTWNQ 6
           ||| |
Db       2 mtwgq 6


RESULT  15
AAY23130
ID   AAY23130 standard; peptide; 6 AA.
XX
AC   AAY23130;
XX
DT   23-AUG-1999  (first entry)
XX
DE   Opioid peptide which inhibits binding of enkephalin.
XX
KW   Opioid peptide; ligand binding; opioid receptor;
KW   micro-selective opioid peptide; enkephalin; opioid receptor system;
KW   blocking; peripheral effect; centrally acting pain killer; morphine.
XX
OS   Synthetic.
XX
FH   Key            Location/Qualifiers
FT   Misc-difference 1..6
FT                   /note= "with the exception of Gly in position 5,
FT                          all these residues are in the D-form"
FT   Modified-site   6
FT                   /note= "amidated"
XX
PN   US5919897-A.
XX
PD   06-JUL-1999.
XX
PF   07-JUN-1995;   95US-0488659.
XX
PR   07-JUN-1995;   95US-0488659.
XX
PA   (TORR-) TORREY PINES INST MOLECULAR STUDIES.
XX
PI   Dooley CT,  Houghten RA;
XX
DR   WPI; 1999-394647/33.
XX
PT   New opioid peptides useful for blocking the peripheral effects of
PT   centrally acting pain killers such as morphine
XX
PS   Example 1; Column 11; 92pp; English.
XX
CC   The specification describes opioid peptides, in which each of the
CC   N atoms in the peptide backbone between respective amino acids is
CC   modified by permethylation, perallylation, perethylation, perbenzylation
CC   and pernaphthylation. The peptides inhibit ligand binding to an opioid
CC   receptor. Specifically, the peptides inhibit the micro-selective
CC   opioid peptide enkephalin. The peptides can be used in vivo
CC   diagnostically to localize opioid receptor subtypes. They can be used
CC   to treat pathologies associated with other compounds which interact with
CC   the opioid receptor system. The peptides are especially useful for
CC   blocking the peripheral effects of centrally acting pain killers such
CC   as morphine. AAY23113-Y23132 represent opioid peptides of the invention,
CC   and are derived from the general sequence given in AAY23112.
XX
SQ   Sequence   6 AA;


  Query Match             46.4%;  Score 26;  DB 20;  Length 6;
  Best Local Similarity   80.0%;  Pred. No. 3.4e+05;
  Matches    4;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 MTWNQ 6
           ||| |
Db       2 mtwgq 6


Search completed: October 23, 2001, 13:28:56
Job time: 404 sec
```

```
OS   Homo sapiens.
XX
PN   WO9946284-A2.
XX
PD   16-SEP-1999.
XX
PF   10-MAR-1999;   99WO-US05284.
XX
PR   13-MAR-1998;   98US-0042107.
PR   26-FEB-1999;   99US-0042107.
XX
PA   (BURN-) BURNHAM INST.
XX
PI   Rajotte D,  Pasqualini R,  Ruoslahti EI;
XX
DR   WPI; 1999-571717/48.
XX
PT   New peptides which selectively home to organs or tissues, used for,
PT   e.g. identifying target ligands and for therapy of pathological
PT   conditions  -
XX
PS   Claim 11; Page 144; 193pp; English.
XX
CC   The present invention describes peptides that selectively home to a
CC   tissue or organ. The peptides can be used for identifying an organ
CC   or tissue, for identifying a target molecule expressed by an organ or
CC   tissue or for treating an organ or tissue pathology, where the organ or
CC   tissue is selected from prostate, lung, skin, retina, pancreas, gut,
CC   ovary, adrenal gland, liver, and lymph node. The peptide bind to the
CC   membrane dipeptidase (MDP). AAY48618 to AAY49066 represent sequences
CC   which are used in the exemplification of the present invention.
XX
SQ   Sequence   9 AA;


  Query Match             50.0%;  Score 28;  DB 20;  Length 9;
  Best Local Similarity   37.5%;  Pred. No. 3.4e+05;
  Matches    3;  Conservative    4;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        1 CMTWNQMN 8
              ||:|: ::
Db        1 cmswdavs 8


RESULT  13
AAW29017
ID   AAW29017 standard; peptide; 6 AA.
XX
AC   AAW29017;
XX
DT   20-JAN-1998  (first entry)
XX
DE   Opioid peptide.
XX
KW   enkephalin; mu-opioid receptor ligand; agonist; antagonist.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 1
FT                   /note= "D-form residue"
FT   Misc-difference 2
FT                   /note= "D-form residue"
FT   Misc-difference 3
FT                   /note= "D-form residue"
FT   Misc-difference 4
FT                   /note= "D-form residue"
FT   Misc-difference 6
FT                   /note= "D-form residue"
FT   Modified-site   6
FT                   /note= "the C-terminal is in amide form"
XX
PN   US5641861-A.
```

```
XX
PD   24-JUN-1997.
XX
PF   07-JUN-1995;   95US-0487006.
XX
PR   07-JUN-1995;   95US-0487006.
XX
PA   (TORR-) TORREY PINES INST MOLECULAR STUDIES.
XX
PI   Dooley CT,  Houghten RA;
XX
DR   WPI; 1997-340994/31.
XX
PT   New opioid peptides which bind mu receptors specifically - have
PT   agonist or antagonist activity and are used for study and
PT   localisation of mu receptors and to treat peripheral side effects of
PT   morphine etc.
XX
PS   Disclosure; Column 12; 92pp; English.
XX
CC   The patent discloses the following new peptides, which are opioids which
CC   bind specifically to the mu receptor: Ac-Phe-Arg-Trp-Trp-Tyr-Xaa-NH2 (1);
CC   Ac-Arg-Trp-Ile-Gly-Trp-Xaa-NH2 (2); Trp-Trp-Pro-Lys-His-Xaa-NH2 (3);
CC   Trp-Trp-Pro-Xaa1-NH2 (4); Tyr-Pro-Phe-Gly-Phe-Xaa-NH2 (5);
CC   D-Ile-D-Met-D-Ser-D-Trp-D-Trp-(Gly)n-Xaa2-NH2 (6);
CC   D-Ile-D-Met-D-Thr-D-Trp-Gly-Xaa2-NH2 (7); Tyr-A1-B2-C3-NH2 (214);
CC   Pm and red ((Me)x(H)y-Tyr-(NMe)z-Tyr-(Xaa3)z-NH2) (221); and
CC   Trp-Trp-Pro-D4-(His)z-(Xaa)z-NH2 (222); where Xaa = any natural amino
CC   acid; Xaa1 = Lys or Arg; n and z = 0 or 1; Xaa2 = Gly or the D form of
CC   any naturally occurring amino acid; A1 = D-norvaline or D-norleucine;
CC   B2 = Gly, Phe or Trp; C3 = Trp or naphthylalanine; x and y = 0-2, but
CC   not over 2 in total; Xaa3 = Phe, DPhe or benzylamino; D4 = Lys or Arg;
CC   Pm and red indicate permethylation and reduction of all CO in peptide
CC   links to methylene. These new compounds are useful: (i) for in vitro
CC   assay and study of opiate receptor subtypes, particularly mu receptors
CC   in the brain; (ii) for in vivo localisation of receptor subtypes; and
CC   (iii) therapeutically to block the peripheral effects (e.g. constipation
CC   and pruritus) of centrally acting pain killers such as morphine.
CC   They are very selective for the mu opioid receptor, over binding to the
CC   delta and kappa receptor subtypes.
CC   The present sequence is a specific example of a peptide (7).
XX
SQ   Sequence   6 AA;


  Query Match             46.4%;  Score 26;  DB 18;  Length 6;
  Best Local Similarity   80.0%;  Pred. No. 3.4e+05;
  Matches    4;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        2 MTWNQ 6
              |||  |
Db        2 mtwgq 6


RESULT  14
AAW24378
ID   AAW24378 standard; peptide; 6 AA.
XX
AC   AAW24378;
XX
DT   25-SEP-1997  (first entry)
XX
DE   New peptide which acts as mu-opioid receptor ligand.
XX
KW   mu-receptor; opioid; opiate; agonist; antagonist; diagnosis;
KW   analgesic.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 1
FT                   /note= "D-form residue"
FT   Misc-difference 2
```

```
XX
DE   T cell epitope/MHC ligand SEQ ID NO:324.
XX
KW   Cytotoxic T-lymphocyte response; CTL; antigen; lymphatic system;
KW   immunisation; tumour; infectious disease; immunotherapy; cancer;
KW   malignant melanoma; viral disease; hepatitis; AIDS.
XX
OS   Synthetic.
OS   Mouse hepatitis virus.
XX
PN   WO9902183-A2.
XX
PD   21-JAN-1999.
XX
PF   10-JUL-1998;   98WO-US14289.
XX
PR   10-DEC-1997;   97US-0988320.
PR   10-JUL-1997;   97CA-2209815.
XX
PA   (CTLI-) CTL IMMUNOTHERAPIES CORP.
XX
PI   Kuendig TM,  Simard JJL;
XX
DR   WPI; 1999-120514/10.
XX
PT   Inducing a cytotoxic T lymphocyte response - by maintaining a level
PT   of antigen in the lymphatic system of a mammal so as to provide a
PT   sustained CTL response, used to treat, e.g. AIDS
XX
PS   Disclosure; Page 36; 199pp; English.
XX
CC   The present invention describes a method of inducing and/or sustaining
CC   an immunological cytotoxic T lymphocyte (CTL) response in a mammal. The
CC   method comprises: (a) delivering an antigen to the mammal at a level to
CC   induce an immunological CTL response in the mammal; and (b) maintaining
CC   the level of the antigen in the mammal's lymphatic system to maintain
CC   the immunologic CTL response. The method can be used for the delivery of
CC   e.g. a differentiation antigen, a tumour-specific multilineage antigen,
CC   an embryonic antigen, an oncogene antigen, a mutated tumour-suppressor
CC   gene antigen, or a viral antigen. They can be used for the treatment of
CC   disease such as cancer, e.g. malignant melanoma or infectious disease,
CC   e.g. viral disease such as hepatitis or AIDS. Sustained antigen delivery
CC   to the lymphatic system provides for potent CTL stimulation that takes
CC   place in the milieu of the lymphoid organ, and it sustains stimulation
CC   that is necessary to keep CTL active, cytotoxic and recirculating
CC   through the body. AAY10071 to AAY10639 represent examples of peptide
CC   antigens given in the present invention.
XX
SQ   Sequence   9 AA;


  Query Match             53.6%;  Score 30;  DB 20;  Length 9;
  Best Local Similarity   44.4%;  Pred. No. 3.4e+05;
  Matches    4;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy        1 CMTWNQMNL 9
            |::||  :|
Db        1 clswngphl 9

RESULT 11
AAY98719
ID   AAY98719 standard; Peptide; 9 AA.
XX
AC   AAY98719;
XX
DT   31-JUL-2000  (first entry)
XX
DE   WT1 derived immunogenic peptide SEQ ID NO:234.
XX
KW   WT1; immunotherapy; immunogenic; malignant disease; cancer; leukaemia;
KW   metastatic disease; mouse; human; Wilm's tumour; immune response;
KW   vaccine.
```

```
XX
OS   Homo sapiens.
XX
PN   WO200018795-A2.
XX
PD   06-APR-2000.
XX
PF   30-SEP-1999;   99WO-US22819.
XX
PR   30-SEP-1998;   98US-0164223.
PR   25-MAR-1999;   99US-0276484.
XX
PA   (CORI-) CORIXA CORP.
PA   (GAIG/) GAIGER A.
XX
PI   Gaiger A,  Cheever M;
XX
DR   WPI; 2000-293107/25.
XX
PT   Novel polypeptides comprising an immunogenic portion of a native WT1
PT   polypeptide, useful for inhibiting the development of malignant
PT   diseases associated with WT1 expression e.g. leukemia or cancer  -
XX
PS   Claim 4; Page 178; 193pp; English.
XX
CC   The present invention describes polypeptides (I) comprising an
CC   immunogenic portion of a native Wilm's Tumour gene product polypeptide,
CC   WT1, (or variants of the immunogenic portion retaining the ability to
CC   react with WT1-specific antisera and/or T-cell lines or clones) and
CC   comprising 16 consecutive amino acids (aa) or less of a native WT1
CC   polypeptide. The polypeptides are useful therapeutically and to
CC   manufacture medicaments for enhancing/inducing an immune response in
CC   patients. The polypeptides, mimetics or polynucleotides can be included
CC   with a carrier/excipient in pharmaceutical compositions or with a
CC   non-specific immune response enhancer (e.g. an adjuvant or enhancer
CC   preferentially enhancing a T cell response) in vaccines. Pharmaceutical
CC   compositions and vaccines can be administered to human patients to
CC   enhance or induce an immune response specific for WT1 or a cell
CC   expressing WT1, useful to inhibit the development of malignant diseases
CC   associated with WT1 expression, e.g. leukemia (especially acute/chronic
CC   myeloid leukemia or acute lymphocytic leukemia) or cancer (especially
CC   breast, lung, thyroid or gastrointestinal cancer, or a melanoma).
CC   AAY98501 to AAY98811 represent polypeptide sequences, and AAA13848 to
CC   AAA13862 represent PCR primers, used in the exemplification of the
CC   present invention.
XX
SQ   Sequence   9 AA;

  Query Match             53.6%;  Score 30;  DB 21;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 3.4e+05;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 CMTW 4
            ||||
Db        6 cmtw 9

RESULT 12
AAY48652
ID   AAY48652 standard; Peptide; 9 AA.
XX
AC   AAY48652;
XX
DT   10-DEC-1999  (first entry)
XX
DE   Membrane dipeptidase-binding lung homing peptide #23.
XX
KW   Homing peptide; organ; tissue; lung; pancreas; skin; retina; MDP;
KW   prostate; ovary; lymph node; adrenal gland; liver; gut; tumour;
KW   membrane dipeptidase.
XX
OS   Synthetic.
```

```
XX
AC   AAY98730;
XX
DT   31-JUL-2000  (first entry)
XX
DE   WT1 derived immunogenic peptide SEQ ID NO:245.
XX
KW   WT1; immunotherapy; immunogenic; malignant disease; cancer; leukaemia;
KW   metastatic disease; mouse; human; Wilm's tumour; immune response;
KW   vaccine.
XX
OS   Homo sapiens.
XX
PN   WO200018795-A2.
XX
PD   06-APR-2000.
XX
PF   30-SEP-1999;   99WO-US22819.
XX
PR   30-SEP-1998;   98US-0164223.
PR   25-MAR-1999;   99US-0276484.
XX
PA   (CORI-) CORIXA CORP.
PA   (GAIG/) GAIGER A.
XX
PI   Gaiger A,  Cheever M;
XX
DR   WPI; 2000-293107/25.
XX
PT   Novel polypeptides comprising an immunogenic portion of a native WT1
PT   polypeptide, useful for inhibiting the development of malignant
PT   diseases associated with WT1 expression e.g. leukemia or cancer  -
XX
PS   Claim 4; Page 179; 193pp; English.
XX
CC   The present invention describes polypeptides (I) comprising an
CC   immunogenic portion of a native Wilm's Tumour gene product polypeptide,
CC   WT1, (or variants of the immunogenic portion retaining the ability to
CC   react with WT1-specific antisera and/or T-cell lines or clones) and
CC   comprising 16 consecutive amino acids (aa) or less of a native WT1
CC   polypeptide. The polypeptides are useful therapeutically and to
CC   manufacture medicaments for enhancing/inducing an immune response in
CC   patients. The polypeptides, mimetics or polynucleotides can be included
CC   with a carrier/excipient in pharmaceutical compositions or with a
CC   non-specific immune response enhancer (e.g. an adjuvant or enhancer
CC   preferentially enhancing a T cell response) in vaccines. Pharmaceutical
CC   compositions and vaccines can be administered to human patients to
CC   enhance or induce an immune response specific for WT1 or a cell
CC   expressing WT1, useful to inhibit the development of malignant diseases
CC   associated with WT1 expression, e.g. leukemia (especially acute/chronic
CC   myeloid leukemia or acute lymphocytic leukemia) or cancer (especially
CC   breast, lung, thyroid or gastrointestinal cancer, or a melanoma).
CC   AAY98501 to AAY98811 represent polypeptide sequences, and AAA13848 to
CC   AAA13862 represent PCR primers, used in the exemplification of the
CC   present invention.
XX
SQ   Sequence   9 AA;

  Query Match             66.1%;  Score 37;  DB 21;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 3.4e+05;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       4 WNQMNL 9
           ||||||
Db       1 wnqmnl 6

RESULT   9
AAY98790
ID   AAY98790 standard; Peptide; 9 AA.
XX
AC   AAY98790;
XX
DT   31-JUL-2000  (first entry)
XX
DE   WT1 derived immunogenic peptide SEQ ID NO:305.
XX
KW   WT1; immunotherapy; immunogenic; malignant disease; cancer; leukaemia;
KW   metastatic disease; mouse; human; Wilm's tumour; immune response;
KW   vaccine.
XX
OS   Mus musculus.
XX
PN   WO200018795-A2.
XX
PD   06-APR-2000.
XX
PF   30-SEP-1999;   99WO-US22819.
XX
PR   30-SEP-1998;   98US-0164223.
PR   25-MAR-1999;   99US-0276484.
XX
PA   (CORI-) CORIXA CORP.
PA   (GAIG/) GAIGER A.
XX
PI   Gaiger A,  Cheever M;
XX
DR   WPI; 2000-293107/25.
XX
PT   Novel polypeptides comprising an immunogenic portion of a native WT1
PT   polypeptide, useful for inhibiting the development of malignant
PT   diseases associated with WT1 expression e.g. leukemia or cancer  -
XX
PS   Claim 4; Page 188; 193pp; English.
XX
CC   The present invention describes polypeptides (I) comprising an
CC   immunogenic portion of a native Wilm's Tumour gene product polypeptide,
CC   WT1, (or variants of the immunogenic portion retaining the ability to
CC   react with WT1-specific antisera and/or T-cell lines or clones) and
CC   comprising 16 consecutive amino acids (aa) or less of a native WT1
CC   polypeptide. The polypeptides are useful therapeutically and to
CC   manufacture medicaments for enhancing/inducing an immune response in
CC   patients. The polypeptides, mimetics or polynucleotides can be included
CC   with a carrier/excipient in pharmaceutical compositions or with a
CC   non-specific immune response enhancer (e.g. an adjuvant or enhancer
CC   preferentially enhancing a T cell response) in vaccines. Pharmaceutical
CC   compositions and vaccines can be administered to human patients to
CC   enhance or induce an immune response specific for WT1 or a cell
CC   expressing WT1, useful to inhibit the development of malignant diseases
CC   associated with WT1 expression, e.g. leukemia (especially acute/chronic
CC   myeloid leukemia or acute lymphocytic leukemia) or cancer (especially
CC   breast, lung, thyroid or gastrointestinal cancer, or a melanoma).
CC   AAY98501 to AAY98811 represent polypeptide sequences, and AAA13848 to
CC   AAA13862 represent PCR primers, used in the exemplification of the
CC   present invention.
XX
SQ   Sequence   9 AA;

  Query Match             66.1%;  Score 37;  DB 21;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 3.4e+05;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       4 WNQMNL 9
           ||||||
Db       1 wnqmnl 6

RESULT  10
AAY10394
ID   AAY10394 standard; Peptide; 9 AA.
XX
AC   AAY10394;
XX
DT   12-MAY-1999  (first entry)
```

```
RESULT   6
AAY98763
ID   AAY98763 standard; Peptide; 9 AA.
XX
AC   AAY98763;
XX
DT   31-JUL-2000  (first entry)
XX
DE   WT1 derived immunogenic peptide SEQ ID NO:278.
XX
KW   WT1; immunotherapy; immunogenic; malignant disease; cancer; leukaemia;
KW   metastatic disease; mouse; human; Wilm's tumour; immune response;
KW   vaccine.
XX
OS   Mus musculus.
XX
PN   WO200018795-A2.
XX
PD   06-APR-2000.
XX
PF   30-SEP-1999;   99WO-US22819.
XX
PR   30-SEP-1998;   98US-0164223.
PR   25-MAR-1999;   99US-0276484.
XX
PA   (CORI-) CORIXA CORP.
PA   (GAIG/) GAIGER A.
XX
PI   Gaiger A,  Cheever M;
XX
DR   WPI; 2000-293107/25.
XX
PT   Novel polypeptides comprising an immunogenic portion of a native WT1
PT   polypeptide, useful for inhibiting the development of malignant
PT   diseases associated with WT1 expression e.g. leukemia or cancer  -
XX
PS   Claim 4; Page 184; 193pp; English.
XX
CC   The present invention describes polypeptides (I) comprising an
CC   immunogenic portion of a native Wilm's Tumour gene product polypeptide,
CC   WT1, (or variants of the immunogenic portion retaining the ability to
CC   react with WT1-specific antisera and/or T-cell lines or clones) and
CC   comprising 16 consecutive amino acids (aa) or less of a native WT1
CC   polypeptide. The polypeptides are useful therapeutically and to
CC   manufacture medicaments for enhancing/inducing an immune response in
CC   patients. The polypeptides, mimetics or polynucleotides can be included
CC   with a carrier/excipient in pharmaceutical compositions or with a
CC   non-specific immune response enhancer (e.g. an adjuvant or enhancer
CC   preferentially enhancing a T cell response) in vaccines. Pharmaceutical
CC   compositions and vaccines can be administered to human patients to
CC   enhance or induce an immune response specific for WT1 or a cell
CC   expressing WT1, useful to inhibit the development of malignant diseases
CC   associated with WT1 expression, e.g. leukemia (especially acute/chronic
CC   myeloid leukemia or acute lymphocytic leukemia) or cancer (especially
CC   breast, lung, thyroid or gastrointestinal cancer, or a melanoma).
CC   AAY98501 to AAY98811 represent polypeptide sequences, and AAA13848 to
CC   AAA13862 represent PCR primers, used in the exemplification of the
CC   present invention.
XX
SQ   Sequence   9 AA;

  Query Match             82.1%;  Score 46;  DB 21;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 3.4e+05;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 CMTWNQM 7
          |||||||
Db      3 cmtwnqm 9

RESULT   7
```

```
AAY98652
ID   AAY98652 standard; Peptide; 9 AA.
XX
AC   AAY98652;
XX
DT   31-JUL-2000  (first entry)
XX
DE   WT1 derived immunogenic peptide SEQ ID NO:167.
XX
KW   WT1; immunotherapy; immunogenic; malignant disease; cancer; leukaemia;
KW   metastatic disease; mouse; human; Wilm's tumour; immune response;
KW   vaccine.
XX
OS   Homo sapiens.
XX
PN   WO200018795-A2.
XX
PD   06-APR-2000.
XX
PF   30-SEP-1999;   99WO-US22819.
XX
PR   30-SEP-1998;   98US-0164223.
PR   25-MAR-1999;   99US-0276484.
XX
PA   (CORI-) CORIXA CORP.
PA   (GAIG/) GAIGER A.
XX
PI   Gaiger A,  Cheever M;
XX
DR   WPI; 2000-293107/25.
XX
PT   Novel polypeptides comprising an immunogenic portion of a native WT1
PT   polypeptide, useful for inhibiting the development of malignant
PT   diseases associated with WT1 expression e.g. leukemia or cancer  -
XX
PS   Claim 4; Page 168; 193pp; English.
XX
CC   The present invention describes polypeptides (I) comprising an
CC   immunogenic portion of a native Wilm's Tumour gene product polypeptide,
CC   WT1, (or variants of the immunogenic portion retaining the ability to
CC   react with WT1-specific antisera and/or T-cell lines or clones) and
CC   comprising 16 consecutive amino acids (aa) or less of a native WT1
CC   polypeptide. The polypeptides are useful therapeutically and to
CC   manufacture medicaments for enhancing/inducing an immune response in
CC   patients. The polypeptides, mimetics or polynucleotides can be included
CC   with a carrier/excipient in pharmaceutical compositions or with a
CC   non-specific immune response enhancer (e.g. an adjuvant or enhancer
CC   preferentially enhancing a T cell response) in vaccines. Pharmaceutical
CC   compositions and vaccines can be administered to human patients to
CC   enhance or induce an immune response specific for WT1 or a cell
CC   expressing WT1, useful to inhibit the development of malignant diseases
CC   associated with WT1 expression, e.g. leukemia (especially acute/chronic
CC   myeloid leukemia or acute lymphocytic leukemia) or cancer (especially
CC   breast, lung, thyroid or gastrointestinal cancer, or a melanoma).
CC   AAY98501 to AAY98811 represent polypeptide sequences, and AAA13848 to
CC   AAA13862 represent PCR primers, used in the exemplification of the
CC   present invention.
XX
SQ   Sequence   9 AA;

  Query Match             73.2%;  Score 41;  DB 21;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 3.4e+05;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 CMTWNQ 6
          ||||||
Db      4 cmtwnq 9

RESULT   8
AAY98730
ID   AAY98730 standard; Peptide; 9 AA.
```

```
CC   myeloid leukemia or acute lymphocytic leukemia) or cancer (especially
CC   breast, lung, thyroid or gastrointestinal cancer, or a melanoma).
CC   AAY98501 to AAY98811 represent polypeptide sequences, and AAA13848 to
CC   AAA13862 represent PCR primers, used in the exemplification of the
CC   present invention.
XX
SQ   Sequence   9 AA;


  Query Match             100.0%;  Score 56;  DB 21;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 3.4e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CMTWNQMNL 9
             |||||||||
Db       1 cmtwnqmnl 9


RESULT   4
AAY80202
ID   AAY80202 standard; Peptide; 9 AA.
XX
AC   AAY80202;
XX
DT   24-MAY-2000  (first entry)
XX
DE   Human Wilms' tumour suppressor gene WT1 product peptide SEQ ID NO:7.
XX
KW   Wilm's tumour suppressor gene; WT1; cancer; antigen; vaccine; MHC;
KW   major histocompatibility complex; leukaemia; tumour; antitumour.
XX
OS   Homo sapiens.
XX
PN   WO200006602-A1.
XX
PD   10-FEB-2000.
XX
PF   30-JUL-1999;   99WO-JP04130.
XX
PR   31-JUL-1998;   98JP-0218093.
XX
PA   (SUGI/) SUGIYAMA H.
XX
PI   Sugiyama H,  Oka Y;
XX
DR   WPI; 2000-195264/17.
XX
PT   Cancer antigens based on Wilm's tumor suppressor gene WT1 product or
PT   peptide derivatives, for cancer vaccines in treating leukemia and solid
PT   tumors e.g. stomach cancer, skin cancer, lung cancer and breast cancer
XX
PS   Claim 5; Page 18; 48pp; Japanese.
XX
CC   The present invention describes a cancer antigen containing the active
CC   component of Wilms' tumour suppressor gene WT1 product, or partial
CC   peptides, for cancer vaccines in treating leukaemia and solid tumours.
CC   The cancer antigens are useful for cancer vaccines in treating
CC   leukaemia, bone-marrow abnormal formation syndrome, malignant lymphoma,
CC   multiple myeloma, stomach cancer, cancer of the large intestine, lung
CC   cancer, breast cancer, blastoma, liver cancer, skin cancer, bladder
CC   cancer, prostate cancer, uterus cancer, cervical cancer, or ovary
CC   cancer. The present sequence represents a peptide from the human
CC   Wilms' tumour suppressor gene WT1 product.
XX
SQ   Sequence   9 AA;


  Query Match             100.0%;  Score 56;  DB 21;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 3.4e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CMTWNQMNL 9
             |||||||||
Db       1 cmtwnqmnl 9


RESULT   5
AAY98616
ID   AAY98616 standard; Peptide; 9 AA.
XX
AC   AAY98616;
XX
DT   31-JUL-2000  (first entry)
XX
DE   WT1 derived immunogenic peptide SEQ ID NO:131.
XX
KW   WT1; immunotherapy; immunogenic; malignant disease; cancer; leukaemia;
KW   metastatic disease; mouse; human; Wilm's tumour; immune response;
KW   vaccine.
XX
OS   Homo sapiens.
XX
PN   WO200018795-A2.
XX
PD   06-APR-2000.
XX
PF   30-SEP-1999;   99WO-US22819.
XX
PR   30-SEP-1998;   98US-0164223.
PR   25-MAR-1999;   99US-0276484.
XX
PA   (CORI-) CORIXA CORP.
PA   (GAIG/) GAIGER A.
XX
PI   Gaiger A,  Cheever M;
XX
DR   WPI; 2000-293107/25.
XX
PT   Novel polypeptides comprising an immunogenic portion of a native WT1
PT   polypeptide, useful for inhibiting the development of malignant
PT   diseases associated with WT1 expression e.g. leukemia or cancer  -
XX
PS   Claim 4; Page 163; 193pp; English.
XX
CC   The present invention describes polypeptides (I) comprising an
CC   immunogenic portion of a native Wilm's Tumour gene product polypeptide,
CC   WT1, (or variants of the immunogenic portion retaining the ability to
CC   react with WT1-specific antisera and/or T-cell lines or clones) and
CC   comprising 16 consecutive amino acids (aa) or less of a native WT1
CC   polypeptide. The polypeptides are useful therapeutically and to
CC   manufacture medicaments for enhancing/inducing an immune response in
CC   patients. The polypeptides, mimetics or polynucleotides can be included
CC   with a carrier/excipient in pharmaceutical compositions or with a
CC   non-specific immune response enhancer (e.g. an adjuvant or enhancer
CC   preferentially enhancing a T cell response) in vaccines. Pharmaceutical
CC   compositions and vaccines can be administered to human patients to
CC   enhance or induce an immune response specific for WT1 or a cell
CC   expressing WT1, useful to inhibit the development of malignant diseases
CC   associated with WT1 expression, e.g. leukemia (especially acute/chronic
CC   myeloid leukemia or acute lymphocytic leukemia) or cancer (especially
CC   breast, lung, thyroid or gastrointestinal cancer, or a melanoma).
CC   AAY98501 to AAY98811 represent polypeptide sequences, and AAA13848 to
CC   AAA13862 represent PCR primers, used in the exemplification of the
CC   present invention.
XX
SQ   Sequence   9 AA;


  Query Match             82.1%;  Score 46;  DB 21;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 3.4e+05;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CMTWNQM 7
             |||||||
Db       3 cmtwnqm 9
```

```
CC   The present sequence is peptide epitope WT235-43, produced by WT1
CC   expressing cells and found at residues 235-243 of the WT1 protein, which
CC   is recognised by cytotoxic T lymphocytes. WT1 is aberrantly expressed in
CC   leukaemias, breast cancer, melanoma and ovarian cancer. The peptide can
CC   be used as a vaccine to stimulate the elimination, by cytotoxic T
CC   lymphocytes, of cancer cells aberrantly expressing WT1. In addition, the
CC   nucleic acid encoding the peptide may also be used in the same manner.
CC   Alternatively, the peptide may be used in vitro to produce activated
CC   cytotoxic T lymphocytes.
XX
SQ   Sequence   9 AA;

  Query Match             100.0%;  Score 56;  DB 21;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 3.4e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         1 CMTWNQMNL 9
               |||||||||
Db         1 cmtwnqmnl 9


RESULT   2
AAY98534
ID   AAY98534 standard; Peptide; 9 AA.
XX
AC   AAY98534;
XX
DT   31-JUL-2000  (first entry)
XX
DE   WT1 derived immunogenic peptide SEQ ID NO:49.
XX
KW   WT1; immunotherapy; immunogenic; malignant disease; cancer; leukaemia;
KW   metastatic disease; mouse; human; Wilm's tumour; immune response;
KW   vaccine.
XX
OS   Homo sapiens.
XX
PN   WO200018795-A2.
XX
PD   06-APR-2000.
XX
PF   30-SEP-1999;   99WO-US22819.
XX
PR   30-SEP-1998;   98US-0164223.
PR   25-MAR-1999;   99US-0276484.
XX
PA   (CORI-) CORIXA CORP.
PA   (GAIG/) GAIGER A.
XX
PI   Gaiger A,  Cheever M;
XX
DR   WPI; 2000-293107/25.
XX
PT   Novel polypeptides comprising an immunogenic portion of a native WT1
PT   polypeptide, useful for inhibiting the development of malignant
PT   diseases associated with WT1 expression e.g. leukemia or cancer  -
XX
PS   Claim 4; Page 151; 193pp; English.
XX
CC   The present invention describes polypeptides (I) comprising an
CC   immunogenic portion of a native Wilm's Tumour gene product polypeptide,
CC   WT1, (or variants of the immunogenic portion retaining the ability to
CC   react with WT1-specific antisera and/or T-cell lines or clones) and
CC   comprising 16 consecutive amino acids (aa) or less of a native WT1
CC   polypeptide. The polypeptides are useful therapeutically and to
CC   manufacture medicaments for enhancing/inducing an immune response in
CC   patients. The polypeptides, mimetics or polynucleotides can be included
CC   with a carrier/excipient in pharmaceutical compositions or with a
CC   non-specific immune response enhancer (e.g. an adjuvant or enhancer
CC   preferentially enhancing a T cell response) in vaccines. Pharmaceutical
CC   compositions and vaccines can be administered to human patients to
CC   enhance or induce an immune response specific for WT1 or a cell
```

```
CC   expressing WT1, useful to inhibit the development of malignant diseases
CC   associated with WT1 expression, e.g. leukemia (especially acute/chronic
CC   myeloid leukemia or acute lymphocytic leukemia) or cancer (especially
CC   breast, lung, thyroid or gastrointestinal cancer, or a melanoma).
CC   AAY98501 to AAY98811 represent polypeptide sequences, and AAA13848 to
CC   AAA13862 represent PCR primers, used in the exemplification of the
CC   present invention.
XX
SQ   Sequence   9 AA;

  Query Match             100.0%;  Score 56;  DB 21;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 3.4e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         1 CMTWNQMNL 9
               |||||||||
Db         1 cmtwnqmnl 9


RESULT   3
AAY98743
ID   AAY98743 standard; Peptide; 9 AA.
XX
AC   AAY98743;
XX
DT   31-JUL-2000  (first entry)
XX
DE   WT1 derived immunogenic peptide SEQ ID NO:258.
XX
KW   WT1; immunotherapy; immunogenic; malignant disease; cancer; leukaemia;
KW   metastatic disease; mouse; human; Wilm's tumour; immune response;
KW   vaccine.
XX
OS   Mus musculus.
XX
PN   WO200018795-A2.
XX
PD   06-APR-2000.
XX
PF   30-SEP-1999;   99WO-US22819.
XX
PR   30-SEP-1998;   98US-0164223.
PR   25-MAR-1999;   99US-0276484.
XX
PA   (CORI-) CORIXA CORP.
PA   (GAIG/) GAIGER A.
XX
PI   Gaiger A,  Cheever M;
XX
DR   WPI; 2000-293107/25.
XX
PT   Novel polypeptides comprising an immunogenic portion of a native WT1
PT   polypeptide, useful for inhibiting the development of malignant
PT   diseases associated with WT1 expression e.g. leukemia or cancer  -
XX
PS   Claim 4; Page 181; 193pp; English.
XX
CC   The present invention describes polypeptides (I) comprising an
CC   immunogenic portion of a native Wilm's Tumour gene product polypeptide,
CC   WT1, (or variants of the immunogenic portion retaining the ability to
CC   react with WT1-specific antisera and/or T-cell lines or clones) and
CC   comprising 16 consecutive amino acids (aa) or less of a native WT1
CC   polypeptide. The polypeptides are useful therapeutically and to
CC   manufacture medicaments for enhancing/inducing an immune response in
CC   patients. The polypeptides, mimetics or polynucleotides can be included
CC   with a carrier/excipient in pharmaceutical compositions or with a
CC   non-specific immune response enhancer (e.g. an adjuvant or enhancer
CC   preferentially enhancing a T cell response) in vaccines. Pharmaceutical
CC   compositions and vaccines can be administered to human patients to
CC   enhance or induce an immune response specific for WT1 or a cell
CC   expressing WT1, useful to inhibit the development of malignant diseases
CC   associated with WT1 expression, e.g. leukemia (especially acute/chronic
```

```
                          GenCore version 4.5
                 Copyright (c) 1993 - 2000  Compugen Ltd.


OM protein - protein search, using sw model

Run on:          October 23, 2001, 13:28:56 ; Search time 31.82 Seconds
                                              (without alignments)
                                              17.147 Million cell updates/sec

Title:           US-09-625-963-2
Perfect score:   56
Sequence:        1 CMTWNQMNL 9

Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched:        412676 seqs, 60623988 residues

Total number of hits satisfying chosen parameters:      75637

Minimum DB seq length: 0
Maximum DB seq length: 9

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       A_Geneseq_0601:*
                1:  /SIDS8/gcgdata/geneseq/geneseqp/AA1980.DAT:*
                2:  /SIDS8/gcgdata/geneseq/geneseqp/AA1981.DAT:*
                3:  /SIDS8/gcgdata/geneseq/geneseqp/AA1982.DAT:*
                4:  /SIDS8/gcgdata/geneseq/geneseqp/AA1983.DAT:*
                5:  /SIDS8/gcgdata/geneseq/geneseqp/AA1984.DAT:*
                6:  /SIDS8/gcgdata/geneseq/geneseqp/AA1985.DAT:*
                7:  /SIDS8/gcgdata/geneseq/geneseqp/AA1986.DAT:*
                8:  /SIDS8/gcgdata/geneseq/geneseqp/AA1987.DAT:*
                9:  /SIDS8/gcgdata/geneseq/geneseqp/AA1988.DAT:*
                10:  /SIDS8/gcgdata/geneseq/geneseqp/AA1989.DAT:*
                11:  /SIDS8/gcgdata/geneseq/geneseqp/AA1990.DAT:*
                12:  /SIDS8/gcgdata/geneseq/geneseqp/AA1991.DAT:*
                13:  /SIDS8/gcgdata/geneseq/geneseqp/AA1992.DAT:*
                14:  /SIDS8/gcgdata/geneseq/geneseqp/AA1993.DAT:*
                15:  /SIDS8/gcgdata/geneseq/geneseqp/AA1994.DAT:*
                16:  /SIDS8/gcgdata/geneseq/geneseqp/AA1995.DAT:*
                17:  /SIDS8/gcgdata/geneseq/geneseqp/AA1996.DAT:*
                18:  /SIDS8/gcgdata/geneseq/geneseqp/AA1997.DAT:*
                19:  /SIDS8/gcgdata/geneseq/geneseqp/AA1998.DAT:*
                20:  /SIDS8/gcgdata/geneseq/geneseqp/AA1999.DAT:*
                21:  /SIDS8/gcgdata/geneseq/geneseqp/AA2000.DAT:*
                22:  /SIDS8/gcgdata/geneseq/geneseqp/AA2001.DAT:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                                SUMMARIES
                    %
Result            Query
  No.    Score   Match Length  DB  ID                      Description
 ------------------------------------------------------------------------
    1      56   100.0      9   21  AAY94203                 Human cytotoxic T
    2      56   100.0      9   21  AAY98534                 WT1 derived immuno
    3      56   100.0      9   21  AAY98743                 WT1 derived immuno
    4      56   100.0      9   21  AAY80202                 Human Wilms' tumou
    5      46    82.1      9   21  AAY98616                 WT1 derived immuno
    6      46    82.1      9   21  AAY98763                 WT1 derived immuno
    7      41    73.2      9   21  AAY98652                 WT1 derived immuno
    8      37    66.1      9   21  AAY98730                 WT1 derived immuno
    9      37    66.1      9   21  AAY98790                 WT1 derived immuno
   10      30    53.6      9   20  AAY10394                 T cell epitope/MHC
   11      30    53.6      9   21  AAY98719                 WT1 derived immuno
   12      28    50.0      9   20  AAY48652                 Membrane dipeptida
   13      26    46.4      6   18  AAW29017                 Opioid peptide.  S
   14      26    46.4      6   18  AAW24378                 New peptide which
   15      26    46.4      6   20  AAY23130                 Opioid peptide whi
   16      26    46.4      9   21  AAY98636                 WT1 derived immuno
   17      26    46.4      9   21  AAY98771                 WT1 derived immuno
   18      25    44.6      6   17  AAR86220                 Anti-ELAM-1 bindin
   19      25    44.6      6   19  AAW64039                 ELAM-1 peptide mim
   20      25    44.6      8   20  AAY14272                 Chemokine peptide,
   21      24    42.9      6   17  AAR86217                 Anti-ELAM-1 bindin
   22      24    42.9      6   19  AAW64036                 ELAM-1 peptide mim
   23      24    42.9      8   15  AAR50842                 Cyclic peptide inh
   24      24    42.9      8   15  AAR50843                 Cyclic peptide inh
   25      24    42.9      8   22  AAB90154                 Factor VIIa (FVIIa
   26      24    42.9      8   22  AAB90156                 Factor VIIa (FVIIa
   27      24    42.9      8   22  AAB90158                 Factor VIIa (FVIIa
   28      24    42.9      9   15  AAR50820                 Cyclic peptide inh
   29      24    42.9      9   15  AAR50821                 Cyclic peptide inh
   30      24    42.9      9   15  AAR50822                 Cyclic peptide inh
   31      24    42.9      9   15  AAR50914                 Cyclic peptide inh
   32      24    42.9      9   15  AAR50840                 Cyclic peptide inh
   33      24    42.9      9   15  AAR50841                 Cyclic peptide inh
   34      24    42.9      9   15  AAR50835                 Cyclic peptide inh
   35      24    42.9      9   22  AAB90155                 Factor VIIa (FVIIa
   36      24    42.9      9   22  AAB90157                 Factor VIIa (FVIIa
   37      23    41.1      4   22  AAB49875                 Human endostatin p
   38      23    41.1      5   17  AAR86216                 Anti-ELAM-1 bindin
   39      23    41.1      5   17  AAR86221                 Anti-ELAM-1 bindin
   40      23    41.1      5   17  AAR86222                 Anti-ELAM-1 bindin
   41      23    41.1      5   17  AAR86223                 Anti-ELAM-1 bindin
   42      23    41.1      5   17  AAR86224                 Anti-ELAM-1 bindin
   43      23    41.1      5   17  AAR86225                 Anti-ELAM-1 bindin
   44      23    41.1      5   19  AAW64040                 ELAM-1 peptide mim
   45      23    41.1      5   19  AAW64042                 ELAM-1 peptide mim


                                ALIGNMENTS


RESULT   1
AAY94203
ID   AAY94203 standard; peptide; 9 AA.
XX
AC   AAY94203;
XX
DT   28-JUL-2000  (first entry)
XX
DE   Human cytotoxic T lymphocyte-recognised WT1 peptide WT235-43.
XX
KW   WT235-43; peptide; epitope; Wilm's tumour gene; leukaemia;
KW   breast cancer; melanoma; ovarian cancer; immunotherapy.
XX
OS   Homo sapiens.
XX
PN   WO200026249-A1.
XX
PD   11-MAY-2000.
XX
PF   02-NOV-1999;   99WO-GB03572.
XX
PR   02-NOV-1998;   98GB-0023897.
XX
PA   (IMCO-) IMPERIAL COLLEGE INNOVATIONS LTD.
XX
PI   Stauss HJ,  Gao L;
XX
DR   WPI; 2000-376123/32.
XX
PT   Novel peptides comprising WT-1 and GATA-1 epitopes, their fragments or
PT   variants, useful as vaccines for cancer immunotherapy  -
XX
PS   Claim 2; Page 74; 93pp; English.
XX
```

```
Qy       3 TW 4
           ||
Db       7 TW 8


RESULT  15
Q69100
ID   Q69100      PRELIMINARY;      PRT;     9 AA.
AC   Q69100;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   PROTEIN 32 (FRAGMENT).
OS   Herpes simplex virus (type 2).
OC   Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC   Alphaherpesvirinae; Simplexvirus.
OX   NCBI_TaxID=10310;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90156509; PubMed=2154597;
RA   Yei S., Chowdhury S.I., Bhat B.M., Conley A.J., Wold W.S.,
RA   Batterson W.;
RT   "Identification and characterization of the herpes simplex virus type
RT   2 gene encoding the essential capsid protein ICP32/VP19c.";
RL   J. Virol. 64:1124-1134(1990).
DR   EMBL; M33905; AAA45847.1; -.
FT   NON_TER       9      9
SQ   SEQUENCE   9 AA;  993 MW;  C13BCDC5BB1DC37D CRC64;

  Query Match             28.6%;  Score 16;  DB 14;  Length 9;
  Best Local Similarity   66.7%;  Pred. No. 4.2e+05;
  Matches    2;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 MTW 4
           | |
Db       1 MAW 3


Search completed: October 23, 2001, 13:33:46
Job time: 288 sec
```

```
  Matches       3;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       5 NQMNL 9
           | :||
Db       5 NMVNL 9


RESULT  11
Q43928
ID   Q43928      PRELIMINARY;      PRT;     9 AA.
AC   Q43928; Q43918; Q43920; Q43921;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   DNA FOR EXEF-EXEG INTERGENIC REGION (FRAGMENT).
GN   EXEF'.
OS   Aeromonas caviae (Aeromonas formicans).
OC   Bacteria; Proteobacteria; gamma subdivision; Aeromonadaceae;
OC   Aeromonas.
OX   NCBI_TaxID=643;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=VARIOUS STRAINS;
RX   MEDLINE=97089747; PubMed=8935655;
RA   Karlyshev A.V., Macintyre S.;
RT   "Study of the intergenic exeF-exeG region and its application as a
RT   simple preliminary test for Aeromonas spp.";
RL   FEMS Microbiol. Lett. 137:37-44(1996).
DR   EMBL; X89464; CAA61643.1; -.
DR   EMBL; X89462; CAA61639.1; -.
DR   EMBL; X89460; CAA61635.1; -.
DR   EMBL; X89463; CAA61641.1; -.
FT   NON_TER       1      1
SQ   SEQUENCE   9 AA;  1059 MW;  785852C69444472B CRC64;

  Query Match            30.4%;  Score 17;  DB 2;  Length 9;
  Best Local Similarity  60.0%;  Pred. No. 4.2e+05;
  Matches       3;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       5 NQMNL 9
           | :||
Db       5 NMVNL 9


RESULT  12
Q62721
ID   Q62721      PRELIMINARY;      PRT;     8 AA.
AC   Q62721;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   PROHIBITIN (FRAGMENT).
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=FISHER;
RX   MEDLINE=95331633; PubMed=7607556;
RA   Altus M.S., Wood C.M., Stewart D.A., Roskams A.I., Friedman V.,
RA   Henderson T., Owens G.A., Danner D.B., Jupe E.R., Dell'Orco R.T.,
RA   McClung J.K.;
RT   "Regions of evolutionary conservation between the rat and human
RT   prohibitin-encoding genes.";
RL   Gene 158:291-294(1995).
DR   EMBL; U17178; AAA86692.1; -.
FT   NON_TER       8      8
SQ   SEQUENCE   8 AA;  1150 MW;  EFD3237B05A41376 CRC64;

  Query Match            28.6%;  Score 16;  DB 11;  Length 8;
  Best Local Similarity  66.7%;  Pred. No. 4.2e+05;
  Matches       2;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       4 WNQ 6
           | |
Db       6 WKQ 8


RESULT  13
P79940
ID   P79940      PRELIMINARY;      PRT;     8 AA.
AC   P79940;
DT   01-MAY-1997 (TrEMBLrel. 03, Created)
DT   01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT   01-MAY-1997 (TrEMBLrel. 03, Last annotation update)
DE   XMEIS1-4 PROTEIN (FRAGMENT).
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Steelman S., Moskow J.J., Muzynski K., North C., Druck T.,
RA   Montgomery J.C., Huebner K., Daar I.O., Buchberg A.M.;
RL   Submitted (AUG-1996) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U68389; AAB19199.1; -.
FT   NON_TER       1      1
SQ   SEQUENCE   8 AA;  1187 MW;  278B51F37B11F40B CRC64;

  Query Match            28.6%;  Score 16;  DB 13;  Length 8;
  Best Local Similarity  50.0%;  Pred. No. 4.2e+05;
  Matches       2;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       4 WNQM 7
           |: |
Db       5 WHYM 8


RESULT  14
Q9MW43
ID   Q9MW43      PRELIMINARY;      PRT;     9 AA.
AC   Q9MW43;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   MHC CLASS I ANTIGEN (FRAGMENT).
GN   HLA-B39061.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=BLOOD;
RX   MEDLINE=97246738; PubMed=9089103;
RA   Vargas-Alarcon G., Gomez-Casado E., Martinez-Laso J., Granados J.,
RA   Layrisse Z., Alegre R., Arnaiz-Villena A.;
RT   "Differences in intron 2 sequences between B*39061 and B*39062 in
RT   Amerindians: comparison with those of B*3901, B*5101, and B*52012
RT   alleles.";
RL   Immunogenetics 45:436-439(1997).
DR   EMBL; L76640; AAF86298.1; -.
KW   MHC.
FT   NON_TER       1      1
FT   NON_TER       9      9
SQ   SEQUENCE   9 AA;  975 MW;  51C661AB01A72DD7 CRC64;

  Query Match            28.6%;  Score 16;  DB 7;  Length 9;
  Best Local Similarity  100.0%;  Pred. No. 4.2e+05;
  Matches       2;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
RL   Submitted (MAY-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF064963; AAD09947.1; -.
KW   Hypothetical protein.
FT   NON_TER       1      1
SQ   SEQUENCE   8 AA;  993 MW;  046B5AA453772727 CRC64;


  Query Match             32.1%;  Score 18;  DB 2;  Length 8;
  Best Local Similarity   50.0%;  Pred. No. 4.2e+05;
  Matches    2;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        2 MTWN 5
             :||
Db        2 LLWN 5


RESULT   7
Q85723
ID   Q85723      PRELIMINARY;      PRT;     9 AA.
AC   Q85723;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   SIS ONCOGENE (FRAGMENT).
OS   Simian sarcoma virus.
OC   Viruses; Retroid viruses; Retroviridae; Mammalian type C retroviruses.
OX   NCBI_TaxID=11817;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=84106822; PubMed=6319011;
RA   Devare S.G., Shatzman A.R., Robbins K.C., Rosenberg M., Aaronson S.A.;
RT   "Expression of the PDGF-related transforming protein of simian sarcoma
RT   virus in E. coli.";
RL   Cell 36:43-49(1984).
DR   EMBL; K01473; AAA46816.1; -.
FT   NON_TER       9      9
SQ   SEQUENCE   9 AA;  1048 MW;  9C53A866C361A731 CRC64;


  Query Match             32.1%;  Score 18;  DB 14;  Length 9;
  Best Local Similarity   66.7%;  Pred. No. 4.2e+05;
  Matches    2;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        2 MTW 4
            :||
Db        3 LTW 5


RESULT   8
Q44001
ID   Q44001      PRELIMINARY;      PRT;     9 AA.
AC   Q44001;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   DNA FOR EXEF-EXEG INTERGENIC REGION (STRAIN CDC 9179-79) (FRAGMENT).
GN   EXEF'.
OS   Aeromonas eucrenophila.
OC   Bacteria; Proteobacteria; gamma subdivision; Aeromonadaceae;
OC   Aeromonas.
OX   NCBI_TaxID=649;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CDC 9179-79;
RX   MEDLINE=97089747; PubMed=8935655;
RA   Karlyshev A.V., Macintyre S.;
RT   "Study of the intergenic exeF-exeG region and its application as a
RT   simple preliminary test for Aeromonas spp.";
RL   FEMS Microbiol. Lett. 137:37-44(1996).
DR   EMBL; X89461; CAA61637.1; -.
FT   NON_TER       1      1
SQ   SEQUENCE   9 AA;  1059 MW;  785852C69444472B CRC64;
```

```
  Query Match             30.4%;  Score 17;  DB 2;  Length 9;
  Best Local Similarity   60.0%;  Pred. No. 4.2e+05;
  Matches    3;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        5 NQMNL 9
            | :||
Db        5 NMVNL 9


RESULT   9
Q44377
ID   Q44377      PRELIMINARY;      PRT;     9 AA.
AC   Q44377;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   DNA FOR EXEF-EXEG INTERGENIC REGION (STRAIN ATCC 49659) (FRAGMENT).
GN   EXEF'.
OS   Aeromonas trota.
OC   Bacteria; Proteobacteria; gamma subdivision; Aeromonadaceae;
OC   Aeromonas.
OX   NCBI_TaxID=653;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 49659;
RX   MEDLINE=97089747; PubMed=8935655;
RA   Karlyshev A.V., Macintyre S.;
RT   "Study of the intergenic exeF-exeG region and its application as a
RT   simple preliminary test for Aeromonas spp.";
RL   FEMS Microbiol. Lett. 137:37-44(1996).
DR   EMBL; X89468; CAA61651.1; -.
FT   NON_TER       1      1
SQ   SEQUENCE   9 AA;  1059 MW;  785852C69444472B CRC64;


  Query Match             30.4%;  Score 17;  DB 2;  Length 9;
  Best Local Similarity   60.0%;  Pred. No. 4.2e+05;
  Matches    3;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        5 NQMNL 9
            | :||
Db        5 NMVNL 9


RESULT  10
Q44468
ID   Q44468      PRELIMINARY;      PRT;     9 AA.
AC   Q44468;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   DNA FOR EXEF-EXEG INTERGENIC REGION (STRAIN CDC 1306-83) (FRAGMENT).
GN   EXEF'.
OS   Aeromonas veronii.
OC   Bacteria; Proteobacteria; gamma subdivision; Aeromonadaceae;
OC   Aeromonas.
OX   NCBI_TaxID=654;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CDC 1306-83;
RX   MEDLINE=97089747; PubMed=8935655;
RA   Karlyshev A.V., Macintyre S.;
RT   "Study of the intergenic exeF-exeG region and its application as a
RT   simple preliminary test for Aeromonas spp.";
RL   FEMS Microbiol. Lett. 137:37-44(1996).
DR   EMBL; X89457; CAA61629.1; -.
FT   NON_TER       1      1
SQ   SEQUENCE   9 AA;  1059 MW;  785852C69444472B CRC64;


  Query Match             30.4%;  Score 17;  DB 2;  Length 9;
  Best Local Similarity   60.0%;  Pred. No. 4.2e+05;
```

```
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94061059; PubMed=7694728;
RA   Spritz R.A., Holmes S.A., Berg S.Z., Nordlund J.J., Fukai K.;
RT   "A recurrent deletion in the KIT (mast/stem cell growth factor
RT   receptor) proto-oncogene is a frequent cause of human piebaldism.";
RL   Hum. Mol. Genet. 2:1499-1500(1993).
DR   EMBL; S67686; AAD13996.1; -.
FT   NON_TER       1      1
SQ   SEQUENCE   9 AA;  1182 MW;  0BC504032361B5AB CRC64;

  Query Match             37.5%;  Score 21;  DB 4;  Length 9;
  Best Local Similarity   50.0%;  Pred. No. 4.2e+05;
  Matches    3;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      3 TWNQMN 8
          || : |
Db      4 TWKREN 9

RESULT   3
Q35792
ID   Q35792      PRELIMINARY;      PRT;     8 AA.
AC   Q35792;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   OXI3 LOCUS IN YEAST MITOCHONDRION (STRAIN D273-10B) (CODES FOR
DE   CYTOCHROME OXIDASE SUBUNIT 1).
OS   Saccharomyces cerevisiae (Baker's yeast).
OG   Mitochondrion.
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=81069885; PubMed=6254986;
RA   Bonitz S.G., Coruzzi G., Thalenfeld B.E., Tzagoloff A., Macino G.;
RT   "Assembly of the mitochondrial membrane system. Structure and
RT   nucleotide sequence of the gene coding for subunit 1 of yeast
RT   cytochrme oxidase.";
RL   J. Biol. Chem. 255:11927-11941(1980).
DR   EMBL; V00694; CAA24063.1; -.
KW   Mitochondrion.
SQ   SEQUENCE   8 AA;  1025 MW;  084693345B5AA337 CRC64;

  Query Match             35.7%;  Score 20;  DB 8;  Length 8;
  Best Local Similarity   60.0%;  Pred. No. 4.2e+05;
  Matches    3;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      4 WNQMN 8
          :|:||
Db      4 YNKMN 8

RESULT   4
Q9UCQ9
ID   Q9UCQ9      PRELIMINARY;      PRT;     9 AA.
AC   Q9UCQ9;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   AUTOTAXIN.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=92129337; PubMed=1733949;
RA   Stracke M.L., Krutzsch H.C., Unsworth E.J., Arestad A., Cioce V.,
RA   Schiffmann E., Liotta L.A.;
RT   "Identification, purification, and partial sequence analysis of
RT   autotaxin, a novel motility-stimulating protein.";
RL   J. Biol. Chem. 267:2524-2529(1992).
SQ   SEQUENCE   9 AA;  1136 MW;  9A3CAB14536772CA CRC64;

  Query Match             33.9%;  Score 19;  DB 4;  Length 9;
  Best Local Similarity   66.7%;  Pred. No. 4.2e+05;
  Matches    2;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      4 WNQ 6
          ||:
Db      5 WNE 7

RESULT   5
P82003
ID   P82003      PRELIMINARY;      PRT;     9 AA.
AC   P82003;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   PROTHORACICOSTATIC PEPTIDE (PTSP).
OS   Bombyx mori (Silk moth).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
OC   Bombycoidea; Bombycidae; Bombyx.
OX   NCBI_TaxID=7091;
RN   [1]
RP   SEQUENCE.
RC   STRAIN=C145 X N140; TISSUE=BRAIN;
RA   Hua Y.-J., Tanaka Y., Nakamura K.;
RT   "Identification of a prothoracicostatic peptide (PTSP) from the larval
RT   brain of the silkworm, Bombyx mori.";
RL   J. Biol. Chem. 0:0-0(1999).
CC   -!- FUNCTION: INHIBITS ECDYSTEROIDOGENESIS BY PROTHORACIC GLAND IN THE
CC       SILKWORM.
CC   -!- DEVELOPMENTAL STAGE: EARLY FIFTH INSTAR.
KW   Amidation.
FT   MOD_RES       9      9       AMIDATION.
SQ   SEQUENCE   9 AA;  1090 MW;  3878C5B4472AB6C3 CRC64;

  Query Match             33.9%;  Score 19;  DB 5;  Length 9;
  Best Local Similarity   40.0%;  Pred. No. 4.2e+05;
  Matches    2;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      4 WNQMN 8
          |  :|
Db      2 WQDLN 6

RESULT   6
O85406
ID   O85406      PRELIMINARY;      PRT;     8 AA.
AC   O85406;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   HYPOTHETICAL 1.0 KDA PROTEIN (FRAGMENT).
OS   Coxiella burnetii.
OC   Bacteria; Proteobacteria; gamma subdivision; Legionellaceae group;
OC   Coxiella group; Coxiella.
OX   NCBI_TaxID=777;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NINE MILE PHASE I;
RA   Willems H., Jaeger C., Baljer G.;
RT   "Physical and genetic map of the obligate intracellular bacterium
RT   Coxiella burnetii.";
```

```
                    GenCore version 4.5
             Copyright (c) 1993 - 2000  Compugen Ltd.

OM protein - protein search, using sw model

Run on:          October 23, 2001, 13:33:46 ; Search time 32.72 Seconds
                                             (without alignments)
                                             36.392 Million cell updates/sec

Title:           US-09-625-963-2
Perfect score:   56
Sequence:        1 CMTWNQMNL 9

Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched:        425026 seqs, 132305027 residues

Total number of hits satisfying chosen parameters:      523

Minimum DB seq length: 0
Maximum DB seq length: 9

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       SPTREMBL_16:*
                1:  sp_archea:*
                2:  sp_bacteria:*
                3:  sp_fungi:*
                4:  sp_human:*
                5:  sp_invertebrate:*
                6:  sp_mammal:*
                7:  sp_mhc:*
                8:  sp_organelle:*
                9:  sp_phage:*
                10:  sp_plant:*
                11:  sp_rodent:*
                12:  sp_unclassified:*
                13:  sp_vertebrate:*
                14:  sp_virus:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                                  SUMMARIES
                 %
Result           Query
  No.   Score    Match Length DB  ID                  Description
---------------------------------------------------------------------------
    1      22    39.3      9  2  Q9R635               Q9r635 chlamydia t
    2      21    37.5      9  4  Q9UMA0               Q9uma0 homo sapien
    3      20    35.7      8  8  Q35792               Q35792 saccharomyc
    4      19    33.9      9  4  Q9UCQ9               Q9ucq9 homo sapien
    5      19    33.9      9  5  P82003               P82003 bombyx mori
    6      18    32.1      8  2  O85406               O85406 coxiella bu
    7      18    32.1      9  14  Q85723               Q85723 simian sarc
    8      17    30.4      9  2  Q44001               Q44001 aeromonas e
    9      17    30.4      9  2  Q44377               Q44377 aeromonas t
   10      17    30.4      9  2  Q44468               Q44468 aeromonas v
   11      17    30.4      9  2  Q43928               Q43928 aeromonas c
   12      16    28.6      8  11  Q62721               Q62721 rattus norv
   13      16    28.6      8  13  P79940               P79940 xenopus lae
   14      16    28.6      9  7  Q9MW43               Q9mw43 homo sapien
   15      16    28.6      9  14  Q69100               Q69100 herpes simp
   16      15    26.8      8  4  Q9UMH9               Q9umh9 homo sapien
   17      15    26.8      9  2  Q57328               Q57328 aeromonas s
   18      15    26.8      9  9  Q38340               Q38340 lactococcus
   19      14    25.0      9  2  O30790               O30790 erwinia amy
   20      14    25.0      9  2  Q9R9C4               Q9r9c4 borrelia bu
   21      14    25.0      9  4  Q16386               Q16386 homo sapien
   22      13    23.2      7  2  O34028               O34028 pseudomonas
   23      13    23.2      8  4  Q15888               Q15888 homo sapien
   24      13    23.2      8  4  Q9UDZ4               Q9udz4 homo sapien
   25      13    23.2      8  11  Q62933               Q62933 rattus norv
   26      13    23.2      8  11  P70243               P70243 mus musculu
   27      13    23.2      8  11  Q9QVD3               Q9qvd3 rattus sp.
   28      13    23.2      9  2  Q9JNI6               Q9jni6 streptococc
   29      13    23.2      9  4  Q9UQW0               Q9uqw0 homo sapien
   30      13    23.2      9  4  Q9UK44               Q9uk44 homo sapien
   31      13    23.2      9  6  Q9GK05               Q9gk05 capra hircu
   32      13    23.2      9  14  Q89491               Q89491 murine minu
   33      12    21.4      7  10  O49223               O49223 glycine max
   34      12    21.4      7  13  O42564               O42564 fugu rubrip
   35      12    21.4      8  5  P82685               P82685 periplaneta
   36      12    21.4      8  5  P82686               P82686 periplaneta
   37      12    21.4      8  5  P82687               P82687 periplaneta
   38      12    21.4      8  5  P82688               P82688 periplaneta
   39      12    21.4      8  5  P82689               P82689 periplaneta
   40      12    21.4      8  11  P82598               P82598 rattus norv
   41      12    21.4      9  8  Q31653               Q31653 anser caeru
   42      12    21.4      9  9  Q38366               Q38366 bacteriopha
   43      12    21.4      9  13  Q9PRJ4               Q9prj4 lepisosteus
   44      12    21.4      9  14  O90350               O90350 hepatitis g
   45      11    19.6      6  13  P82096               P82096 litoria rub

                                  ALIGNMENTS


RESULT   1
Q9R635
ID   Q9R635      PRELIMINARY;      PRT;      9 AA.
AC   Q9R635;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   MAJOR OUTER MEMBRANE PROTEIN VARIABLE DOMAIN IV, MOMP VD IV.
OS   Chlamydia trachomatis.
OC   Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX   NCBI_TaxID=813;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92040090; PubMed=1718870;
RA   Peterson E.M., Cheng X., Markoff B.A., Fielder T.J., de la Maza L.M.;
RT   "Functional and structural mapping of Chlamydia trachomatis species-
RT   specific major outer membrane protein epitopes by use of neutralizing
RT   monoclonal antibodies.";
RL   Infect. Immun. 59:4147-4153(1991).
SQ   SEQUENCE   9 AA;  976 MW;  9C61B041B7645361 CRC64;


  Query Match            39.3%;  Score 22;  DB 2;  Length 9;
  Best Local Similarity  100.0%;  Pred. No. 4.2e+05;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        3 TWN 5
            |||
Db        2 TWN 4


RESULT   2
Q9UMA0
ID   Q9UMA0      PRELIMINARY;      PRT;      9 AA.
AC   Q9UMA0;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   KIT PROTEIN (FRAGMENT).
GN   KIT.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
KW   Neuropeptide; Amidation.
FT   MOD_RES       6      6       AMIDATION.
SQ   SEQUENCE   6 AA;  654 MW;  686365A5B9CDB000 CRC64;


  Query Match             21.4%;  Score 12;  DB 1;  Length 6;
  Best Local Similarity   50.0%;  Pred. No. 9.3e+04;
  Matches    1;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        3 TW 4
            :|
Db        4 SW 5


Search completed: October 23, 2001, 13:34:05
Job time: 277 sec
```

```
CC   -!- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS
CC       (ALPHA, BETA, & GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.
CC   -!- MISCELLANEOUS: CONVERSION OF FIBRINOGEN TO FIBRIN IS TRIGGERED BY
CC       THROMBIN, WHICH CLEAVES FIBRINOPEPTIDES A AND B FROM ALPHA & BETA
CC       CHAINS, AND THUS EXPOSES THE N-TERMINAL POLYMERIZATION SITES
CC       RESPONSIBLE FOR THE FORMATION OF THE SOFT CLOT.
DR   PIR; E28854; E28854.
DR   InterPro; IPR002181; -.
DR   PROSITE; PS00514; FIBRIN_AG_C_DOMAIN; PARTIAL.
KW   Blood coagulation; Plasma.
FT   NON_TER       9      9
SQ   SEQUENCE   9 AA;  1057 MW;  DDFE71E9C7287B06 CRC64;


  Query Match             23.2%;  Score 13;  DB 1;  Length 9;
  Best Local Similarity   60.0%;  Pred. No. 9.3e+04;
  Matches    3;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          5 NQMNL 9
              ||  |
Db          1 NQEGL 5


RESULT 13
FIBB_THEGE
ID   FIBB_THEGE     STANDARD;      PRT;     9 AA.
AC   P19342;
DT   01-NOV-1990 (Rel. 16, Created)
DT   01-NOV-1990 (Rel. 16, Last sequence update)
DT   01-NOV-1990 (Rel. 16, Last annotation update)
DE   FIBRINOPEPTIDE B.
OS   Theropithecus gelada (Gelada baboon).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC   Cercopithecinae; Theropithecus.
OX   NCBI_TaxID=9565;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=84161822; PubMed=6423621;
RA   Nakamura S., Takenaka O., Takahashi K.;
RT   "Fibrinopeptides A and B of baboons (Papio anubis, Papio hamadryas,
RT   and Theropithecus gelada): their amino acid sequences and
RT   evolutionary rates and a molecular phylogeny for the baboons.";
RL   J. Biochem. 94:1973-1978(1983).
CC   -!- FUNCTION: FIBRINOGEN HAS A DOUBLE FUNCTION: YIELDING MONOMERS THAT
CC       POLYMERIZE INTO FIBRIN AND ACTING AS A COFACTOR IN PLATELET
CC       AGGREGATION.
CC   -!- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS
CC       (ALPHA, BETA, & GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.
CC   -!- MISCELLANEOUS: CONVERSION OF FIBRINOGEN TO FIBRIN IS TRIGGERED BY
CC       THROMBIN, WHICH CLEAVES FIBRINOPEPTIDES A AND B FROM ALPHA & BETA
CC       CHAINS, AND THUS EXPOSES THE N-TERMINAL POLYMERIZATION SITES
CC       RESPONSIBLE FOR THE FORMATION OF THE SOFT CLOT.
DR   PIR; F28854; F28854.
DR   InterPro; IPR002181; -.
DR   PROSITE; PS00514; FIBRIN_AG_C_DOMAIN; PARTIAL.
KW   Blood coagulation; Plasma.
FT   NON_TER       9      9
SQ   SEQUENCE   9 AA;  977 MW;  DDFE7879C7287B06 CRC64;


  Query Match             23.2%;  Score 13;  DB 1;  Length 9;
  Best Local Similarity   60.0%;  Pred. No. 9.3e+04;
  Matches    3;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          5 NQMNL 9
              ||  |
Db          1 NQEGL 5


RESULT 14
NEF_HV1Z8
ID   NEF_HV1Z8      STANDARD;      PRT;     9 AA.
AC   P12481;
DT   01-OCT-1989 (Rel. 12, Created)
DT   01-OCT-1989 (Rel. 12, Last sequence update)
DT   01-JUL-1993 (Rel. 26, Last annotation update)
DE   NEGATIVE FACTOR (F-PROTEIN) (27 KDA PROTEIN) (3'ORF) (FRAGMENT).
GN   NEF.
OS   Human immunodeficiency virus type 1 (Z-84 isolate) (HIV-1).
OC   Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX   NCBI_TaxID=11681;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=88281278; PubMed=3395517;
RA   Yourno J., Josephs S.F., Reitz M.S. Jr., Zagury D., Wong-Staal F.,
RA   Gallo R.C.;
RT   "Nucleotide sequence analysis of the env gene of a new Zairian
RT   isolate of HIV-1.";
RL   AIDS Res. Hum. Retroviruses 4:165-173(1988).
CC   -!- FUNCTION: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING
CC       ACTIVITIES, IT SEEM TO DOWN REGULATE THE CD4(T4) ANTIGEN.
CC   -!- MISCELLANEOUS: THE Z-84 ISOLATE WAS TAKEN FROM A 54 YEAR-OLD
CC       ZAIREAN MALE.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; J03653; AAA44687.1; -.
DR   HIV; J03653; NEF$JY1.
KW   AIDS; Myristate; GTP-binding.
FT   LIPID         2      2       MYRISTATE (BY SIMILARITY).
FT   NON_TER       9      9
SQ   SEQUENCE   9 AA;  967 MW;  319CB325A3733878 CRC64;


  Query Match             23.2%;  Score 13;  DB 1;  Length 9;
  Best Local Similarity   33.3%;  Pred. No. 9.3e+04;
  Matches    1;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          4 WNQ 6
              |::
Db          5 WSK 7


RESULT 15
LOK1_LOCMI
ID   LOK1_LOCMI     STANDARD;      PRT;     6 AA.
AC   P41491;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   LOCUSTAKININ I.
OS   Locusta migratoria (Migratory locust).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Orthopteroidea; Orthoptera; Caelifera;
OC   Acridomorpha; Acridoidea; Acrididae; Locusta.
OX   NCBI_TaxID=7004;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Corpora cardiaca;
RX   MEDLINE=92262851; PubMed=1585017;
RA   Schoofs L., Holman G.M., Proost P., van Damme J., Hayes T.K.,
RA   de Loof A.;
RT   "Locustakinin, a novel myotropic peptide from Locusta migratoria,
RT   isolation, primary structure and synthesis.";
RL   Regul. Pept. 37:49-57(1992).
CC   -!- FUNCTION: MYOTROPIC PEPTIDE. MAY BE IMPORTANT IN THE STIMULATION
CC       OF ION TRANSPORT AND INHIBITION OF DIURETIC ACTIVITY IN MALPIGHIAN
CC       TUBULES.
DR   PIR; A61068; A61068.
```

```
OC   Cucujiformia; Tenebrionidae; Tenebrio.
OX   NCBI_TaxID=7067, 7075;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Corpora cardiaca;
RX   MEDLINE=90341081; PubMed=2381871;
RA   Gaede G., Rosinski G.;
RT   "The primary structure of the hypertrehalosemic neuropeptide from
RT   tenebrionid beetles: a novel member of the AKH/RPCH family.";
RL   Peptides 11:455-459(1990).
CC   -!- FUNCTION: HYPERTREHALOSAEMIC FACTORS ARE NEUROPEPTIDES THAT
CC       ELEVATE THE LEVEL OF TREHALOSE IN THE HEMOLYMPH (TREHALOSE IS
CC       THE MAJOR CARBOHYDRATE IN THE HEMOLYMPH OF INSECTS).
CC   -!- SIMILARITY: BELONGS TO THE AKH / HRTH / RPCH FAMILY.
DR   PIR; A43976; A43976.
DR   PIR; B43976; B43976.
DR   InterPro; IPR002047; -.
DR   PROSITE; PS00256; AKH; 1.
KW   Neuropeptide; Amidation.
FT   MOD_RES       1      1       PYRROLIDONE CARBOXYLIC ACID.
FT   MOD_RES       8      8       AMIDATION.
SQ   SEQUENCE   8 AA;  1005 MW;  86745775B9C44736 CRC64;


  Query Match             23.2%;  Score 13;  DB 1;  Length 8;
  Best Local Similarity   66.7%;  Pred. No. 9.3e+04;
  Matches    2;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          6 QMN 8
              |:|
Db          1 QLN 3


RESULT 10
RPCH_PANBO
ID   RPCH_PANBO     STANDARD;      PRT;     8 AA.
AC   P08939;
DT   01-NOV-1988 (Rel. 09, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   RED PIGMENT CONCENTRATING HORMONE (RPCH).
OS   Pandalus borealis (Northern red shrimp).
OC   Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC   Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Caridea; Pandaloidea;
OC   Pandalidae; Pandalus.
OX   NCBI_TaxID=6703;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=75054965; PubMed=4433569;
RA   Fernlund P.;
RT   "Structure of the red-pigment-concentrating hormone of the shrimp,
RT   Pandalus borealis.";
RL   Biochim. Biophys. Acta 371:304-311(1974).
CC   -!- FUNCTION: THIS HORMONE ADAPTS THE ANIMAL TO LIGHT BACKGROUNDS BY
CC       STIMULATING CONCENTRATION OF THE PIGMENT OF ITS RED BODY-
CC       CHROMATOPHORES.
CC   -!- SIMILARITY: BELONGS TO THE AKH / HRTH / RPCH FAMILY.
DR   PIR; S07139; S07139.
DR   InterPro; IPR002047; -.
DR   PROSITE; PS00256; AKH; 1.
KW   Pigment; Hormone; Amidation.
FT   MOD_RES       1      1       PYRROLIDONE CARBOXYLIC ACID.
FT   MOD_RES       8      8       AMIDATION.
SQ   SEQUENCE   8 AA;  948 MW;  86786775B9C44736 CRC64;


  Query Match             23.2%;  Score 13;  DB 1;  Length 8;
  Best Local Similarity   66.7%;  Pred. No. 9.3e+04;
  Matches    2;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          6 QMN 8
              |:|
Db          1 QLN 3
```

```
RESULT 11
FIBB_PAPAN
ID   FIBB_PAPAN     STANDARD;      PRT;     9 AA.
AC   P19344;
DT   01-NOV-1990 (Rel. 16, Created)
DT   01-NOV-1990 (Rel. 16, Last sequence update)
DT   01-NOV-1990 (Rel. 16, Last annotation update)
DE   FIBRINOPEPTIDE B.
OS   Papio anubis (Olive baboon).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC   Cercopithecinae; Papio.
OX   NCBI_TaxID=9555;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=84161822; PubMed=6423621;
RA   Nakamura S., Takenaka O., Takahashi K.;
RT   "Fibrinopeptides A and B of baboons (Papio anubis, Papio hamadryas,
RT   and Theropithecus gelada): their amino acid sequences and
RT   evolutionary rates and a molecular phylogeny for the baboons.";
RL   J. Biochem. 94:1973-1978(1983).
CC   -!- FUNCTION: FIBRINOGEN HAS A DOUBLE FUNCTION: YIELDING MONOMERS THAT
CC       POLYMERIZE INTO FIBRIN AND ACTING AS A COFACTOR IN PLATELET
CC       AGGREGATION.
CC   -!- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS
CC       (ALPHA, BETA, & GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.
CC   -!- MISCELLANEOUS: CONVERSION OF FIBRINOGEN TO FIBRIN IS TRIGGERED BY
CC       THROMBIN, WHICH CLEAVES FIBRINOPEPTIDES A AND B FROM ALPHA & BETA
CC       CHAINS, AND THUS EXPOSES THE N-TERMINAL POLYMERIZATION SITES
CC       RESPONSIBLE FOR THE FORMATION OF THE SOFT CLOT.
DR   PIR; D28854; D28854.
DR   InterPro; IPR002181; -.
DR   PROSITE; PS00514; FIBRIN_AG_C_DOMAIN; PARTIAL.
KW   Blood coagulation; Plasma.
FT   NON_TER       9      9
SQ   SEQUENCE   9 AA;  1076 MW;  DDFE6409C7287B06 CRC64;


  Query Match             23.2%;  Score 13;  DB 1;  Length 9;
  Best Local Similarity   60.0%;  Pred. No. 9.3e+04;
  Matches    3;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          5 NQMNL 9
              ||  |
Db          1 NQEGL 5


RESULT 12
FIBB_PAPHA
ID   FIBB_PAPHA     STANDARD;      PRT;     9 AA.
AC   P19343;
DT   01-NOV-1990 (Rel. 16, Created)
DT   01-NOV-1990 (Rel. 16, Last sequence update)
DT   01-NOV-1990 (Rel. 16, Last annotation update)
DE   FIBRINOPEPTIDE B.
OS   Papio hamadryas (Hamadryas baboon).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC   Cercopithecinae; Papio.
OX   NCBI_TaxID=9557;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=84161822; PubMed=6423621;
RA   Nakamura S., Takenaka O., Takahashi K.;
RT   "Fibrinopeptides A and B of baboons (Papio anubis, Papio hamadryas,
RT   and Theropithecus gelada): their amino acid sequences and
RT   evolutionary rates and a molecular phylogeny for the baboons.";
RL   J. Biochem. 94:1973-1978(1983).
CC   -!- FUNCTION: FIBRINOGEN HAS A DOUBLE FUNCTION: YIELDING MONOMERS THAT
CC       POLYMERIZE INTO FIBRIN AND ACTING AS A COFACTOR IN PLATELET
CC       AGGREGATION.
```

```
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Brain;
RX   MEDLINE=88234141; PubMed=3375140;
RA   Fan Z.W., Eng J., Shaw G., Yalow R.S.;
RT   "Cholecystokinin octapeptide purified from brains of Australian
RT   marsupials.";
RL   Peptides 9:429-431(1988).
CC   -!- FUNCTION: THIS PEPTIDE HORMONE INDUCES GALL BLADDER CONTRACTION
CC       AND THE RELEASE OF PANCREATIC ENZYMES IN THE GUT. ITS FUNCTION
CC       IN THE BRAIN IS NOT CLEAR.
CC   -!- SIMILARITY: BELONGS TO THE GASTRIN/CHOLECYSTOKININ FAMILY.
DR   PIR; A43001; A43001.
DR   PIR; PQ0012; PQ0012.
DR   InterPro; IPR001651; -.
DR   PROSITE; PS00259; GASTRIN; 1.
KW   Amidation; Sulfatation; Hormone.
FT   MOD_RES       2      2       SULFATATION.
FT   MOD_RES       8      8       AMIDATION.
SQ   SEQUENCE   8 AA;  1064 MW;  DDCAA68378768B5A CRC64;


  Query Match             25.0%;  Score 14;  DB 1;  Length 8;
  Best Local Similarity   66.7%;  Pred. No. 9.3e+04;
  Matches    2;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 MTW 4
              | |
Db          3 MGW 5


RESULT   7
DNF1_LOCMI
ID   DNF1_LOCMI     STANDARD;      PRT;     9 AA.
AC   P16339;
DT   01-AUG-1990 (Rel. 15, Created)
DT   01-AUG-1990 (Rel. 15, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   LOCUPRESSIN (DIURETIC NEUROPEPTIDE F1/F2).
OS   Locusta migratoria (Migratory locust).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Orthopteroidea; Orthoptera; Caelifera;
OC   Acridomorpha; Acridoidea; Acrididae; Locusta.
OX   NCBI_TaxID=7004;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Suboesophageal ganglion, and Thoracic ganglion;
RX   MEDLINE=88077077; PubMed=3689410;
RA   Proux J.P., Miller C.A., Li J.P., Carney R.L., Girardie A.,
RA   Delaage M., Schooley D.A.;
RT   "Identification of an arginine vasopressin-like diuretic hormone from
RT   Locusta migratoria.";
RL   Biochem. Biophys. Res. Commun. 149:180-186(1987).
CC   -!- FUNCTION: DIURETIC HORMONE.
CC   -!- SUBUNIT: F2 IS AN ANTIPARALLEL DISULFIDE LINKED DIMER OF F1.
CC   -!- SIMILARITY: BELONGS TO THE VASOPRESSIN/OXYTOCIN FAMILY.
DR   PIR; A29477; A29477.
DR   InterPro; IPR000981; -.
DR   Pfam; PF00220; hormone4; 1.
DR   PROSITE; PS00264; NEUROHYPOPHYS_HORM; 1.
KW   Hormone; Neuropeptide; Amidation.
FT   DISULFID      1      6       IN F1.
FT   DISULFID      1      1       INTERCHAIN (WITH C-6') (IN F2).
FT   DISULFID      6      6       INTERCHAIN (WITH C-1') (IN F2).
FT   MOD_RES       9      9       AMIDATION.
SQ   SEQUENCE   9 AA;  976 MW;  56EB176EB451A057 CRC64;


  Query Match             25.0%;  Score 14;  DB 1;  Length 9;
  Best Local Similarity   40.0%;  Pred. No. 9.3e+04;
  Matches    2;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 CMTWN 5
```

```
                |:  |
Db          1 CLITN 5


RESULT   8
AKH_MELML
ID   AKH_MELML      STANDARD;      PRT;     8 AA.
AC   P25423;
DT   01-MAY-1992 (Rel. 22, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   01-FEB-1994 (Rel. 28, Last annotation update)
DE   ADIPOKINETIC HORMONE (AKH).
OS   Melolontha melolontha (Cockchafer),
OS   Geotrupes stercorosus (Dor beetle), and
OS   Pachnoda marginata (Flower beetle).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;
OC   Scarabaeiformia; Scarabaeidae; Melolonthinae; Melolontha.
OX   NCBI_TaxID=7061, 7087, 7058;
RN   [1]
RP   SEQUENCE.
RC   SPECIES=M.melolontha, and G.stercorosus; TISSUE=Corpora cardiaca;
RX   MEDLINE=91248100; PubMed=2039445;
RA   Gaede G.;
RT   "A unique charged tyrosine-containing member of the adipokinetic
RT   hormone/red-pigment-concentrating hormone peptide family isolated and
RT   sequenced from two beetle species.";
RL   Biochem. J. 275:671-677(1991).
RN   [2]
RP   SEQUENCE.
RC   SPECIES=P.marginata; TISSUE=Corpora cardiaca;
RX   MEDLINE=92265187; PubMed=1586453;
RA   Gaede G., Lopata A., Kellner R., Rinehart K.L. Jr.;
RT   "Primary structures of neuropeptides isolated from the corpora
RT   cardiaca of various cetonid beetle species determined by
RT   pulsed-liquid phase sequencing and tandem fast atom bombardment mass
RT   spectrometry.";
RL   Biol. Chem. Hoppe-Seyler 373:133-142(1992).
CC   -!- FUNCTION: THIS HORMONE, RELEASED FROM CELLS IN THE CORPORA
CC       CARDIACA AFTER THE BEGINNING OF FLIGHT, CAUSES RELEASE OF
CC       DIGLYCERIDES FROM THE FAT BODY AND THEN STIMULATES THE FLIGHT
CC       MUSCLES TO USE THESE DIGLYCERIDES AS AN ENERGY SOURCE.
CC   -!- SIMILARITY: BELONGS TO THE AKH / HRTH / RPCH FAMILY.
DR   PIR; S15422; S15422.
DR   PIR; S21663; S21663.
DR   InterPro; IPR002047; -.
DR   PROSITE; PS00256; AKH; 1.
KW   Neuropeptide; Amidation; Flight.
FT   MOD_RES       1      1       PYRROLIDONE CARBOXYLIC ACID.
FT   MOD_RES       8      8       AMIDATION.
SQ   SEQUENCE   8 AA;  1022 MW;  867AB775AB544736 CRC64;


  Query Match             23.2%;  Score 13;  DB 1;  Length 8;
  Best Local Similarity   66.7%;  Pred. No. 9.3e+04;
  Matches    2;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          6 QMN 8
              |:|
Db          1 QLN 3


RESULT   9
HTF_TENMO
ID   HTF_TENMO      STANDARD;      PRT;     8 AA.
AC   P25419;
DT   01-MAY-1992 (Rel. 22, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   01-FEB-1994 (Rel. 28, Last annotation update)
DE   HYPERTREHALOSAEMIC FACTOR (HOTH) (HYPERTREHALOSEMIC NEUROPEPTIDE).
OS   Tenebrio molitor (Yellow mealworm), and Zophobas rugipes.
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;
```

```
DR   PIR; S33245; S33245.
KW   Neuropeptide; Amidation.
FT   MOD_RES      7      7       AMIDATION.
SQ   SEQUENCE   7 AA;  993 MW;  7362D5B69B041310 CRC64;


  Query Match             35.7%;  Score 20;  DB 1;  Length 7;
  Best Local Similarity   33.3%;  Pred. No. 9.3e+04;
  Matches    2;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      4 WNQMNL 9
          | :|::
Db      1 WREMSV 6


RESULT   3
WWA3_ACHFU
ID   WWA3_ACHFU     STANDARD;      PRT;     7 AA.
AC   P35921;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   01-OCT-1994 (Rel. 30, Last annotation update)
DE   WWAMIDE-3.
OS   Achatina fulica (Giant African snail).
OC   Eukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata; Stylommatophora;
OC   Achatinacea; Achatinidae; Achatina.
OX   NCBI_TaxID=6530;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Ganglion;
RX   MEDLINE=93265912; PubMed=8495720;
RA   Minakata H., Ikeda T., Muneoka Y., Kobayashi M., Nomoto K.;
RT   "WWamide-1, -2 and -3: novel neuromodulatory peptides isolated from
RT   ganglia of the African giant snail, Achatina fulica.";
RL   FEBS Lett. 323:104-108(1993).
DR   PIR; S33244; S33244.
KW   Neuropeptide; Amidation.
FT   MOD_RES      7      7       AMIDATION.
SQ   SEQUENCE   7 AA;  965 MW;  7362D5B69B132310 CRC64;


  Query Match             35.7%;  Score 20;  DB 1;  Length 7;
  Best Local Similarity   33.3%;  Pred. No. 9.3e+04;
  Matches    2;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      4 WNQMNL 9
          | :|::
Db      1 WKEMSV 6


RESULT   4
LMIP_LOCMI
ID   LMIP_LOCMI     STANDARD;      PRT;     9 AA.
AC   P31799;
DT   01-JUL-1993 (Rel. 26, Created)
DT   01-JUL-1993 (Rel. 26, Last sequence update)
DT   01-OCT-1993 (Rel. 27, Last annotation update)
DE   LOCUSTAMYOINHIBITING PEPTIDE (LOM-MIP).
OS   Locusta migratoria (Migratory locust).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Orthopteroidea; Orthoptera; Caelifera;
OC   Acridomorpha; Acridoidea; Acrididae; Locusta.
OX   NCBI_TaxID=7004;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=92179466; PubMed=1796179;
RA   Schoofs L., Holman G.M., Hayes T.K., Nachman R.J., de Loof A.;
RT   "Isolation, identification and synthesis of locustamyoinhibiting
RT   peptide (LOM-MIP), a novel biologically active neuropeptide from
RT   Locusta migratoria.";
RL   Regul. Pept. 36:111-119(1991).
CC   -!- FUNCTION: SUPPRESSES SPONTANEOUS CONTRACTIONS OF THE HINDGUT AND
CC       OVIDUCT.
```

```
CC   -!- TISSUE SPECIFICITY: NEURONS LOCATED IN TWO VENTRAL CELL CLUSTERS
CC       IN THE SUBOESOPHAGEAL GANGLION.
DR   PIR; A60065; AKLQIM.
KW   Amidation; Neuropeptide.
FT   MOD_RES      9      9       AMIDATION.
SQ   SEQUENCE   9 AA;  1060 MW;  387D7DD4472AB6C3 CRC64;


  Query Match             33.9%;  Score 19;  DB 1;  Length 9;
  Best Local Similarity   40.0%;  Pred. No. 9.3e+04;
  Matches    2;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      4 WNQMN 8
          |  :|
Db      2 WQDLN 6


RESULT   5
TMOF_SARBU
ID   TMOF_SARBU     STANDARD;      PRT;     6 AA.
AC   P41495;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   01-FEB-1996 (Rel. 33, Last annotation update)
DE   TRYPSIN-MODULATING OOSTATIC FACTOR (TMOF).
OS   Sarcophaga bullata (Grey flesh fly) (Neobellieria bullata).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Oestroidea; Sarcophagidae; Sarcophaga.
OX   NCBI_TaxID=7385;
RN   [1]
RP   SEQUENCE, AND SYNTHESIS.
RC   TISSUE=Ovary;
RX   MEDLINE=94211930; PubMed=8159807;
RA   Bylemans D., Borovsky D., Hunt D.F., Shabanowitz J., Grauwels L.,
RA   de Loof A.;
RT   "Sequencing and characterization of trypsin modulating oostatic
RT   factor (TMOF) from the ovaries of the grey fleshfly, Neobellieria
RT   (Sarcophaga) bullata.";
RL   Regul. Pept. 50:61-72(1994).
CC   -!- FUNCTION: HAS AN OOSTATIC ACTIVITY. INHIBITS TRYPSIN BIOSYNTHESIS
CC       IN THE MIDGUT WHICH INDIRECTLY REDUCES THE VITELLOGENIN
CC       CONCENTRATION IN THE HEMOLYMPH RESULTING IN INHIBITION OF OOCYTE
CC       DEVELOPMENT.
CC   -!- DEVELOPMENTAL STAGE: SYNTHESIZED AND RELEASED FROM FOLLICULAR
CC       EPITHELIUM AFTER A BLOOD MEAL.
KW   Hormone.
SQ   SEQUENCE   6 AA;  695 MW;  61E72451B7642000 CRC64;


  Query Match             25.0%;  Score 14;  DB 1;  Length 6;
  Best Local Similarity   60.0%;  Pred. No. 9.3e+04;
  Matches    3;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      5 NQMNL 9
          |  ||
Db      1 NPTNL 5


RESULT   6
CCKN_MACEU
ID   CCKN_MACEU     STANDARD;      PRT;     8 AA.
AC   P30369;
DT   01-APR-1993 (Rel. 25, Created)
DT   01-APR-1993 (Rel. 25, Last sequence update)
DT   01-OCT-1996 (Rel. 34, Last annotation update)
DE   CHOLECYSTOKININ (CCK).
GN   CCK.
OS   Macropus eugenii (Tammar wallaby), and
OS   Dasyurus viverrinus (Southeastern quoll).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Metatheria; Diprotodontia; Macropodidae; Macropus.
OX   NCBI_TaxID=9315, 9279;
```

```
                          GenCore version 4.5
                 Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on:          October 23, 2001, 13:34:05 ; Search time 12.6 Seconds
                                              (without alignments)
                                              24.468 Million cell updates/sec

Title:           US-09-625-963-2
Perfect score:   56
Sequence:        1 CMTWNQMNL 9

Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched:        93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters:      213

Minimum DB seq length: 0
Maximum DB seq length: 9

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       SwissProt_39:*

     Pred. No. is the number of results predicted by chance to have a
     score greater than or equal to the score of the result being printed,
     and is derived by analysis of the total score distribution.

                               SUMMARIES
                %
Result        Query
  No.  Score  Match Length DB ID                    Description
-------------------------------------------------------------------------
    1     23   41.1      7  1 WWA2_ACHFU            P35920 achatina fu
    2     20   35.7      7  1 WWA1_ACHFU            P35919 achatina fu
    3     20   35.7      7  1 WWA3_ACHFU            P35921 achatina fu
    4     19   33.9      9  1 LMIP_LOCMI            P31799 locusta mig
    5     14   25.0      6  1 TMOF_SARBU            P41495 sarcophaga
    6     14   25.0      8  1 CCKN_MACEU            P30369 macropus eu
    7     14   25.0      9  1 DNF1_LOCMI            P16339 locusta mig
    8     13   23.2      8  1 AKH_MELML             P25423 melolontha
    9     13   23.2      8  1 HTF_TENMO             P25419 tenebrio mo
   10     13   23.2      8  1 RPCH_PANBO            P08939 pandalus bo
   11     13   23.2      9  1 FIBB_PAPAN            P19344 papio anubi
   12     13   23.2      9  1 FIBB_PAPHA            P19343 papio hamad
   13     13   23.2      9  1 FIBB_THEGE            P19342 theropithec
   14     13   23.2      9  1 NEF_HV128             P12481 human immun
   15     12   21.4      6  1 LOK1_LOCMI            P41491 locusta mig
   16     12   21.4      8  1 AKHG_GRYBI            P14086 gryllus bim
   17     12   21.4      8  1 AKH_LIBAU             P25418 libellula a
   18     12   21.4      8  1 HTF1_PERAM            P04548 periplaneta
   19     12   21.4      8  1 LCK1_LEUMA            P21140 leucophaea
   20     12   21.4      8  1 LCK2_LEUMA            P21141 leucophaea
   21     12   21.4      8  1 LCK3_LEUMA            P21142 leucophaea
   22     12   21.4      8  1 LCK4_LEUMA            P21143 leucophaea
   23     12   21.4      8  1 LCK5_LEUMA            P19987 leucophaea
   24     12   21.4      8  1 LCK6_LEUMA            P19988 leucophaea
   25     12   21.4      8  1 LCK7_LEUMA            P19989 leucophaea
   26     12   21.4      8  1 LCK8_LEUMA            P19990 leucophaea
   27     12   21.4      9  1 CONO_CONST            P05487 conus stria
   28     12   21.4      9  1 OXYT_EISFO            P42998 eisenia foe
   29     12   21.4      9  1 TAL3_PICJA            P17441 pichia jadi
   30     11   19.6      5  1 BPP7_BOTIN            P30425 bothrops in
   31     11   19.6      5  1 UF01_MOUSE            P38639 mus musculu
   32     11   19.6      8  1 ACI_THUAL             P18691 thunnus alb
   33     11   19.6      8  1 AKH_TABAT             P14595 tabanus atr
   34     11   19.6      8  1 HTF2_PERAM            P04549 periplaneta
   35     11   19.6      8  1 PLP_BRANA             P81707 brassica na
   36     11   19.6      8  1 RS10_SALTY            O68928 salmonella
   37     11   19.6      9  1 CONO_CONGE            P05486 conus geogr
   38     11   19.6      9  1 D1_NEPNO              P24816 nephrops no
   39     11   19.6      9  1 DCML_PSECF            P19913 pseudomonas
   40     11   19.6      9  1 DSIP_RABIT            P01158 oryctolagus
   41     11   19.6      9  1 FAR4_CALVO            P41859 calliphora
   42     11   19.6      9  1 FIBB_MACFU            P19345 macaca fusc
   43     11   19.6      9  1 LITO_LITAU            P08945 litoria aur
   44     11   19.6      9  1 LITR_PHYRO            P08946 phyllomedus
   45     11   19.6      9  1 OXYT_CYPCA            P23879 cyprinus ca

                               ALIGNMENTS

RESULT   1
WWA2_ACHFU
ID   WWA2_ACHFU     STANDARD;      PRT;     7 AA.
AC   P35920;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   01-OCT-1994 (Rel. 30, Last annotation update)
DE   WWAMIDE-2.
OS   Achatina fulica (Giant African snail).
OC   Eukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata; Stylommatophora;
OC   Achatinacea; Achatinidae; Achatina.
OX   NCBI_TaxID=6530;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Ganglion;
RX   MEDLINE=93265912; PubMed=8495720;
RA   Minakata H., Ikeda T., Muneoka Y., Kobayashi M., Nomoto K.;
RT   "WWamide-1, -2 and -3: novel neuromodulatory peptides isolated from
RT   ganglia of the African giant snail, Achatina fulica.";
RL   FEBS Lett. 323:104-108(1993).
KW   Neuropeptide; Amidation.
FT   MOD_RES       7      7       AMIDATION.
SQ   SEQUENCE   7 AA;  964 MW;  7362D5B686D32310 CRC64;

  Query Match             41.1%;  Score 23;  DB 1;  Length 7;
  Best Local Similarity   50.0%;  Pred. No. 9.3e+04;
  Matches    3;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       4 WNQMNL 9
           | ||::
Db       1 WKQMSV 6


RESULT   2
WWA1_ACHFU
ID   WWA1_ACHFU     STANDARD;      PRT;     7 AA.
AC   P35919;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   01-OCT-1994 (Rel. 30, Last annotation update)
DE   WWAMIDE-1.
OS   Achatina fulica (Giant African snail).
OC   Eukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata; Stylommatophora;
OC   Achatinacea; Achatinidae; Achatina.
OX   NCBI_TaxID=6530;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Ganglion;
RX   MEDLINE=93265912; PubMed=8495720;
RA   Minakata H., Ikeda T., Muneoka Y., Kobayashi M., Nomoto K.;
RT   "WWamide-1, -2 and -3: novel neuromodulatory peptides isolated from
RT   ganglia of the African giant snail, Achatina fulica.";
RL   FEBS Lett. 323:104-108(1993).
CC   -!- FUNCTION: EXHIBITS MODULATORY EFFECTS ON THE PERIPHERAL NERVOUS
CC       SYSTEM. INHIBITS ACTIVITY ON A CENTRAL NEURON.
```

```
R;Yamada, M.; Wasserman, R.; Reichard, B.A.; Shane, S.; Caton, A.J.; Rovera, G.
J. Exp. Med. 173, 395-407, 1991
A;Title: Preferential utilization of specific immunoglobulin heavy chain diversity and j
A;Reference number: PT0222; MUID:91108337
A;Accession: PT0299
A;Molecule type: DNA
A;Residues: 1-9 <YAM>
A;Experimental source: B lymphocyte
C;Keywords: heterotetramer; immunoglobulin


  Query Match             30.4%;  Score 17;  DB 2;  Length 9;
  Best Local Similarity   33.3%;  Pred. No. 2.2e+05;
  Matches    2;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       3 TWNQMN 8
           :|:  |
Db       3 SWDYSN 8


RESULT  15
S78426
52.5K protein - spiny lobster (fragment)
C;Species: Panulirus argus (spiny lobster)
C;Date: 14-Apr-1998 #sequence_revision 24-Apr-1998 #text_change 19-May-2000
C;Accession: S78426
R;James, M.O.; Boyle, S.M.; Trapido-Rosenthal, H.G.; Smith, W.C.; Greenberg, R.M.; Shive
Arch. Biochem. Biophys. 329, 31-38, 1996
A;Title: cDNA and protein sequence of a major form of P450, CYP2L, in the hepatopancreas
A;Reference number: S68856; MUID:96201120
A;Accession: S78426
A;Molecule type: protein
A;Residues: 1-9 <JAM>
A;Experimental source: hepatopancreas microsomes


  Query Match             28.6%;  Score 16;  DB 2;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.2e+05;
  Matches    2;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3 TW 4
           ||
Db       1 TW 2


Search completed: October 23, 2001, 13:29:50
Job time: 168 sec
```

```
  Best Local Similarity   66.7%;  Pred. No. 2.2e+05;
  Matches    2;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        4 WNQ 6
            ||:
Db        3 WNK 5


RESULT  9
PN0649
alpha-dextrin endo-1,6-alpha-glucosidase (EC 3.2.1.41) - Bacillus sp. (strain S-1) (frag
C;Species: Bacillus sp.
C;Date: 19-May-1994 #sequence_revision 19-May-1994 #text_change 07-May-1999
C;Accession: PN0649
R;Kim, C.H.; Choi, H.I.; Lee, D.S.
Biosci. Biotechnol. Biochem. 57, 1632-1637, 1993
A;Title: Purification and biochemical properties of an alkaline pullulanase from alkalop
A;Reference number: PN0649; MUID:94080025
A;Accession: PN0649
A;Molecule type: protein
A;Residues: 1-7 <KIM>
C;Comment: This enzyme is used togther with glucoamylase to improve the efficiency of st
ntent in high maltose syrups.
C;Keywords: glycosidase; hydrolase; polysaccharide degradation


  Query Match             30.4%;  Score 17;  DB 2;  Length 7;
  Best Local Similarity   66.7%;  Pred. No. 2.2e+05;
  Matches    2;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        2 MTW 4
            |:|
Db        4 MSW 6


RESULT  10
S19288
acylase - Kluyvera cryocrescens
C;Species: Kluyvera cryocrescens
C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 24-Jul-1997
C;Accession: S19288
R;Martin, J.; Slade, A.; Aitken, A.; Arche, R.; Virden, R.
Biochem. J. 280, 659-662, 1991
A;Title: Chemical modification of serine at the active site of penicillin acylase from K
A;Reference number: S19288; MUID:92109664
A;Accession: S19288
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-8 <MAR>


  Query Match             30.4%;  Score 17;  DB 2;  Length 8;
  Best Local Similarity   50.0%;  Pred. No. 2.2e+05;
  Matches    2;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy        1 CMTW 4
            |  |
Db        1 CNMW 4


RESULT  11
A39308
glycine reductase (EC 1.4.99.-) sulfhydryl protein C, alpha chain - Clostridium sticklan
C;Species: Clostridium sticklandii
C;Date: 19-Jun-1992 #sequence_revision 19-Jun-1992 #text_change 15-Aug-1997
C;Accession: A39308
R;Stadtman, T.C.; Davis, J.N.
J. Biol. Chem. 266, 22147-22153, 1991
A;Title: Glycine reductase protein C. Properties and characterization of its role in the
A;Reference number: A39308; MUID:92042141
A;Accession: A39308
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-8 <STA>
C;Function:
A;Description: glycine reductase complex catalyzes the reductive deamination of glyci
C;Keywords: ATP; oxidoreductase


  Query Match             30.4%;  Score 17;  DB 2;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 2.2e+05;
  Matches    2;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        4 WN 5
            ||
Db        6 WN 7


RESULT  12
A41117
acetylcholinesterase (EC 3.1.1.7), venom - Asian cobra (fragment)
C;Species: Naja naja oxiana (Asian cobra, Oxus cobra)
C;Date: 27-Mar-1992 #sequence_revision 27-Mar-1992 #text_change 23-Jun-1993
C;Accession: A41117
R;Kreienkamp, H.J.; Weise, C.; Raba, R.; Aaviksaar, A.; Hucho, F.
Proc. Natl. Acad. Sci. U.S.A. 88, 6117-6121, 1991
A;Title: Anionic subsites of the catalytic center of acetylcholinesterase from Torped
A;Reference number: A41117; MUID:91296772
A;Accession: A41117
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-8 <KRE>
C;Keywords: carboxylic ester hydrolase


  Query Match             30.4%;  Score 17;  DB 2;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 2.2e+05;
  Matches    2;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        4 WN 5
            ||
Db        5 WN 6


RESULT  13
PT0272
Ig heavy chain CRD3 region (clone 3-103B) - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 16-Aug-1996
C;Accession: PT0272
R;Yamada, M.; Wasserman, R.; Reichard, B.A.; Shane, S.; Caton, A.J.; Rovera, G.
J. Exp. Med. 173, 395-407, 1991
A;Title: Preferential utilization of specific immunoglobulin heavy chain diversity an
A;Reference number: PT0222; MUID:91108337
A;Accession: PT0272
A;Molecule type: DNA
A;Residues: 1-9 <YAM>
A;Experimental source: B lymphocyte
C;Keywords: heterotetramer; immunoglobulin


  Query Match             30.4%;  Score 17;  DB 2;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.2e+05;
  Matches    2;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        4 WN 5
            ||
Db        3 WN 4


RESULT  14
PT0299
Ig heavy chain CRD3 region (clone 5-103B) - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 16-Aug-1996
C;Accession: PT0299
```

```
neuromodulatory peptide WWamide-2 - giant African snail
C;Species: Achatina fulica (giant African snail)
C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 24-Jul-1997
C;Accession: S33245
R;Minakata, H.; Ikeda, T.; Muneoka, Y.; Kobayashi, M.; Nomoto, K.
FEBS Lett. 323, 104-108, 1993
A;Title: WWamide-1, -2 and -3: novel neuromodulatory peptides isolated from ganglia of t
A;Reference number: S33244; MUID:93265912
A;Accession: S33245
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-7 <MIN>


  Query Match             35.7%;  Score 20;  DB 2;  Length 7;
  Best Local Similarity   33.3%;  Pred. No. 2.2e+05;
  Matches    2;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       4 WNQMNL 9
           | :|::
Db       1 WREMSV 6

RESULT   4
F41946
T-cell receptor gamma chain (1a.27) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 07-May-1999
C;Accession: F41946
R;Whetsell, M.; Mosley, R.L.; Whetsell, L.; Schaefer, F.V.; Miller, K.S.; Klein, J.R.
Mol. Cell. Biol. 11, 5902-5909, 1991
A;Title: Rearrangement and junctional-site sequence analyses of T-cell receptor gamma ge
A;Reference number: A41946; MUID:92049316
A;Accession: F41946
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-6 <WHE>
C;Keywords: T-cell receptor


  Query Match             33.9%;  Score 19;  DB 2;  Length 6;
  Best Local Similarity   50.0%;  Pred. No. 2.2e+05;
  Matches    2;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 CMTW 4
           |  |
Db       2 CAVW 5

RESULT   5
AKLQIM
locustamyoinhibiting peptide - migratory locust
C;Species: Locusta migratoria (migratory locust)
C;Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 20-Mar-1998
C;Accession: A60065
R;Schoofs, L.; Holman, G.M.; Hayes, T.K.; Nachman, R.J.; De Loof, A.
Regul. Pept. 36, 111-119, 1991
A;Title: Isolation, identification and synthesis of locustamyoinhibiting peptide (LOM-MI
A;Reference number: A60065; MUID:92179466
A;Accession: A60065
A;Molecule type: protein
A;Residues: 1-9 <SCH>
C;Comment: This peptide hormone suppresses spontaneous contractions of the hindgut and c
C;Superfamily: locustamyoinhibiting peptide
C;Keywords: amidated carboxyl end; hormone
F;9/Modified site: amidated carboxyl end (Trp) #status experimental


  Query Match             33.9%;  Score 19;  DB 1;  Length 9;
  Best Local Similarity   40.0%;  Pred. No. 2.2e+05;
  Matches    2;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       4 WNQMN 8
```

```
           |  :|
Db       2 WQDLN 6

RESULT   6
A57444
neuropeptide Grb-AST B1 - two-spotted cricket
C;Species: Gryllus bimaculatus (two-spotted cricket)
C;Date: 26-Jan-1996 #sequence_revision 26-Jan-1996 #text_change 26-Jan-1996
C;Accession: A57444
R;Lorenz, M.W.; Kellner, R.; Hoffmann, K.H.
J. Biol. Chem. 270, 21103-21108, 1995
A;Title: A family of neuropeptides that inhibit juvenile hormone biosynthesis in the
A;Reference number: A57444; MUID:95403341
A;Accession: A57444
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-9 <LOR>


  Query Match             33.9%;  Score 19;  DB 2;  Length 9;
  Best Local Similarity   40.0%;  Pred. No. 2.2e+05;
  Matches    2;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       4 WNQMN 8
           |  :|
Db       2 WQDLN 6

RESULT   7
B57444
neuropeptide Grb-AST B2 - two-spotted cricket
C;Species: Gryllus bimaculatus (two-spotted cricket)
C;Date: 26-Jan-1996 #sequence_revision 26-Jan-1996 #text_change 26-Jan-1996
C;Accession: B57444
R;Lorenz, M.W.; Kellner, R.; Hoffmann, K.H.
J. Biol. Chem. 270, 21103-21108, 1995
A;Title: A family of neuropeptides that inhibit juvenile hormone biosynthesis in the
A;Reference number: A57444; MUID:95403341
A;Accession: B57444
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-9 <LOR>


  Query Match             33.9%;  Score 19;  DB 2;  Length 9;
  Best Local Similarity   40.0%;  Pred. No. 2.2e+05;
  Matches    2;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       4 WNQMN 8
           |  :|
Db       2 WRDLN 6

RESULT   8
JH0253
gut pentapeptide - Japanese eel
C;Species: Anguilla japonica (Japanese eel)
C;Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 11-Apr-1995
C;Accession: JH0253
R;Uesaka, T.; Ikeda, T.; Kubota, I.; Muneoka, Y.; Ando, M.
Biochem. Biophys. Res. Commun. 180, 828-832, 1991
A;Title: Structure and function of a pentapeptide isolated from the gut of the eel.
A;Reference number: JH0253; MUID:92062113
A;Accession: JH0253
A;Molecule type: protein
A;Residues: 1-5 <UES>
A;Experimental source: gut
C;Comment: This peptide increased basal tone of the circular muscle of the esophagoga
, and of the circular muscle of the gastro-intestinal junction.


  Query Match             32.1%;  Score 18;  DB 2;  Length 5;
```

```
                    GenCore version 4.5
             Copyright (c) 1993 - 2000  Compugen Ltd.

OM protein - protein search, using sw model

Run on:          October 23, 2001, 13:29:50 ; Search time 22.25 Seconds
                                              (without alignments)
                                              30.812 Million cell updates/sec

Title:           US-09-625-963-2
Perfect score:   56
Sequence:        1 CMTWNQMNL 9

Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched:        219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters:      787

Minimum DB seq length: 0
Maximum DB seq length: 9

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       PIR_68:*
                1:  pir1:*
                2:  pir2:*
                3:  pir3:*
                4:  pir4:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                              SUMMARIES
                 %
Result          Query
  No.   Score   Match Length DB  ID                          Description
-----------------------------------------------------------------------------
    1     23    41.1      7  2  S33246                     neuromodulatory pe
    2     20    35.7      7  2  S33244                     neuromodulatory pe
    3     20    35.7      7  2  S33245                     neuromodulatory pe
    4     19    33.9      6  2  F41946                     T-cell receptor ga
    5     19    33.9      9  1  AKLQIM                     locustamyoinhibiti
    6     19    33.9      9  2  A57444                     neuropeptide Grb-A
    7     19    33.9      9  2  B57444                     neuropeptide Grb-A
    8     18    32.1      5  2  JH0253                     gut pentapeptide -
    9     17    30.4      7  2  PN0649                     alpha-dextrin endo
   10     17    30.4      8  2  S19288                     acylase - Kluyvera
   11     17    30.4      8  2  A39308                     glycine reductase
   12     17    30.4      8  2  A41117                     acetylcholinestera
   13     17    30.4      9  2  PT0272                     Ig heavy chain CRD
   14     17    30.4      9  2  PT0299                     Ig heavy chain CRD
   15     16    28.6      9  2  S78426                     52.5K protein - sp
   16     14    25.0      6  2  B35640                     cerebellar degener
   17     14    25.0      7  2  B34818                     vicilin 57K chain
   18     14    25.0      8  2  PQ0012                     cholecystokinin -
   19     14    25.0      8  2  A43001                     cholecystokinin -
   20     14    25.0      9  2  G85802                     hypothetical prote
   21     14    25.0      9  2  C57444                     neuropeptide Grb-A
   22     14    25.0      9  2  A29477                     diuretic neuropept
   23     13    23.2      5  2  PT0580                     T-cell receptor be
   24     13    23.2      6  2  A31263                     dihydrofolate redu
   25     13    23.2      7  2  PX0008                     glucuronosyltransf
   26     13    23.2      8  2  A61348                     red pigment-concen
   27     13    23.2      8  2  S15422                     adipokinetic hormo
   28     13    23.2      8  2  A43976                     hypertrehalosemic
   29     13    23.2      8  2  B43976                     hypertrehalosemic
   30     13    23.2      8  2  A58641                     adipokinetic hormo
   31     13    23.2      8  2  A38887                     T-cell receptor ga
   32     13    23.2      9  2  D28854                     fibrinopeptide B -
   33     13    23.2      9  2  E28854                     fibrinopeptide B -
   34     13    23.2      9  2  F28854                     fibrinopeptide B -
   35     13    23.2      9  2  D57444                     neuropeptide Grb-A
   36     12    21.4      4  2  PT0661                     T-cell receptor be
   37     12    21.4      5  2  A60803                     neuropeptide - sea
   38     12    21.4      5  2  PT0281                     Ig heavy chain CRD
   39     12    21.4      6  2  S66195                     alcohol dehydrogen
   40     12    21.4      6  2  B31263                     dihydrofolate redu
   41     12    21.4      6  2  A61068                     locustakinin - mig
   42     12    21.4      6  2  PT0519                     T-cell receptor be
   43     12    21.4      7  2  A61081                     tryptophyllin, bas
   44     12    21.4      7  2  B39127                     phosphotransferase
   45     12    21.4      7  4  I55382                     hypothetical pepti

                              ALIGNMENTS

RESULT   1
S33246
neuromodulatory peptide WWamide-3 - giant African snail
C;Species: Achatina fulica (giant African snail)
C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 24-Jul-1997
C;Accession: S33246
R;Minakata, H.; Ikeda, T.; Muneoka, Y.; Kobayashi, M.; Nomoto, K.
FEBS Lett. 323, 104-108, 1993
A;Title: WWamide-1, -2 and -3: novel neuromodulatory peptides isolated from ganglia o
A;Reference number: S33244; MUID:93265912
A;Accession: S33246
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-7 <MIN>


  Query Match             41.1%;  Score 23;  DB 2;  Length 7;
  Best Local Similarity   50.0%;  Pred. No. 2.2e+05;
  Matches    3;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       4 WNQMNL 9
           | ||::
Db       1 WKQMSV 6


RESULT   2
S33244
neuromodulatory peptide WWamide-1 - giant African snail
C;Species: Achatina fulica (giant African snail)
C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 24-Jul-1997
C;Accession: S33244
R;Minakata, H.; Ikeda, T.; Muneoka, Y.; Kobayashi, M.; Nomoto, K.
FEBS Lett. 323, 104-108, 1993
A;Title: WWamide-1, -2 and -3: novel neuromodulatory peptides isolated from ganglia o
A;Reference number: S33244; MUID:93265912
A;Accession: S33244
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-7 <MIN>


  Query Match             35.7%;  Score 20;  DB 2;  Length 7;
  Best Local Similarity   33.3%;  Pred. No. 2.2e+05;
  Matches    2;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       4 WNQMNL 9
           | :|::
Db       1 WKEMSV 6


RESULT   3
S33245
```

```
J. Biochem. 107, 111-117, 1990
A;Title: Amino acid sequence of glutathione S-transferase b from guinea pig liver.
A;Reference number: JX0095; MUID:90236961
A;Accession: JX0095
A;Molecule type: protein
A;Residues: 1-217 <KAM>
C;Comment: Glutathione transferases are a family of enzymes that are multifunctional in
C;Superfamily: glutathione transferase
C;Keywords: transferase


  Query Match             70.6%;  Score 36;  DB 2;  Length 217;
  Best Local Similarity   85.7%;  Pred. No. 14;
  Matches    6;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      3 FPNAPYL 9
          ||| |||
Db     56 FPNLPYL 62


Search completed: October 23, 2001, 13:22:05
Job time: 118 sec
```

```
                         GenCore version 4.5
               Copyright (c) 1993 - 2000  Compugen Ltd.

OM protein - protein search, using sw model

Run on:         October 23, 2001, 13:20:07 ; Search time 23.71 Seconds
                                             (without alignments)
                                             7.816 Million cell updates/sec

Title:          US-09-625-963-1
Perfect score:  51
Sequence:       1 RMFPNAPYL 9

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched:       197339 seqs, 20590346 residues

Total number of hits satisfying chosen parameters:      197339

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       Issued_Patents_AA:*
               1:  /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
               2:  /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
               3:  /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
               4:  /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
               5:  /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
               6:  /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                                 SUMMARIES
                  %
Result          Query
  No.   Score   Match Length DB  ID                       Description
---------------------------------------------------------------------------
    1      51   100.0    154  1  US-08-102-942A-5          Sequence 5, Appli
    2      51   100.0    210  1  US-08-234-783-2           Sequence 2, Appli
    3      51   100.0    210  1  US-08-456-907-2           Sequence 2, Appli
    4      51   100.0    210  5  PCT-US95-05523-2          Sequence 2, Appli
    5      51   100.0    345  1  US-08-102-942A-2          Sequence 2, Appli
    6      51   100.0    429  1  US-08-234-783-4           Sequence 4, Appli
    7      51   100.0    429  1  US-08-456-907-4           Sequence 4, Appli
    8      51   100.0    429  5  PCT-US95-05523-4          Sequence 4, Appli
    9      51   100.0    449  1  US-08-102-942A-4          Sequence 4, Appli
   10      51   100.0    449  1  US-08-102-942A-6          Sequence 6, Appli
   11      37    72.5     50  4  US-08-975-080-26          Sequence 26, Appl
   12      37    72.5   1151  3  US-08-836-134-23          Sequence 23, Appl
   13      37    72.5   1232  3  US-08-836-134-2           Sequence 2, Appli
   14      36    70.6     67  2  US-08-511-485-23          Sequence 23, Appl
   15      36    70.6    217  4  US-08-297-431B-33         Sequence 33, Appl
   16      36    70.6    370  3  US-09-150-133-1           Sequence 1, Appli
   17      36    70.6    370  3  US-09-150-133-3           Sequence 3, Appli
   18      36    70.6    370  3  US-09-150-141-1           Sequence 1, Appli
   19      36    70.6    370  3  US-09-150-141-3           Sequence 3, Appli
   20      36    70.6    370  4  US-09-374-493-1           Sequence 1, Appli
   21      36    70.6    370  4  US-09-374-493-3           Sequence 3, Appli
   22      36    70.6    370  4  US-09-374-824-1           Sequence 1, Appli
   23      36    70.6    370  4  US-09-374-824-3           Sequence 3, Appli
   24      36    70.6    370  4  US-09-374-492-1           Sequence 1, Appli
   25      36    70.6    370  4  US-09-374-492-3           Sequence 3, Appli
   26      36    70.6    389  4  US-08-430-286A-7          Sequence 7, Appli
   27      36    70.6    438  5  PCT-US95-05922A-2         Sequence 2, Appli
   28      36    70.6    618  2  US-08-511-485-8           Sequence 8, Appli
   29      36    70.6    618  3  US-09-212-971-8           Sequence 8, Appli
   30      36    70.6    618  4  US-08-800-929A-8          Sequence 8, Appli
   31      36    70.6    618  4  US-08-569-749-2           Sequence 2, Appli
   32      36    70.6    618  5  PCT-US96-12860-2          Sequence 2, Appli
   33      35    68.6    380  3  US-09-150-133-9           Sequence 9, Appli
   34      35    68.6    380  3  US-09-150-141-9           Sequence 9, Appli
   35      35    68.6    380  4  US-09-374-493-9           Sequence 9, Appli
   36      35    68.6    380  4  US-09-374-824-9           Sequence 9, Appli
   37      35    68.6    380  4  US-09-374-492-9           Sequence 9, Appli
   38      34    66.7    842  2  US-08-928-692-18          Sequence 18, Appl
   39      33    64.7    267  3  US-08-301-162-16          Sequence 16, Appl
   40      33    64.7    376  3  US-09-150-133-7           Sequence 7, Appli
   41      33    64.7    376  3  US-09-150-141-7           Sequence 7, Appli
   42      33    64.7    376  4  US-09-374-493-7           Sequence 7, Appli
   43      33    64.7    376  4  US-09-374-824-7           Sequence 7, Appli
   44      33    64.7    376  4  US-09-374-492-7           Sequence 7, Appli
   45      33    64.7    377  3  US-09-150-133-5           Sequence 5, Appli

                                 ALIGNMENTS

RESULT   1
US-08-102-942A-5
; Sequence 5, Application US/08102942A
; Patent No. 5726288
;  GENERAL INFORMATION:
;    APPLICANT:  Call, Katherine M.
;    APPLICANT:  Glaser, Thomas M.
;    APPLICANT:  Ito, Caryn Y.
;    APPLICANT:  Buckler, Alan J.
;    APPLICANT:  Pelletier, Jerry
;    APPLICANT:  Haber, Daniel A.
;    APPLICANT:  Rose, Elise A.
;    APPLICANT:  Housman, David E.
;    APPLICANT:  Bruening, Wendy
;    APPLICANT:  Darveau, Andre
;    TITLE OF INVENTION:  Localization and Characterization of the
;    TITLE OF INVENTION:  Wilms' Tumor Gene
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Hamilton, Brook, Smith & Reynolds, P.C.
;      STREET:  Two Militia Drive
;      CITY:  Lexington
;      STATE:  Massachusetts
;      COUNTRY:  U.S.
;      ZIP:  02173
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/102,942A
;      FILING DATE:  02-AUG-1993
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Granahan, Patricia
;      REGISTRATION NUMBER:  32,227
;      REFERENCE/DOCKET NUMBER:  MIT-5194A2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-861-6240
;      TELEFAX:  617-861-9540
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  154 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  not relevant
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-102-942A-5
```

```
  Query Match             100.0%;  Score 51;  DB 1;  Length 154;
  Best Local Similarity   100.0%;  Pred. No. 0.031;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 RMFPNAPYL 9
          |||||||||
Db     40 RMFPNAPYL 48


RESULT   2
US-08-234-783-2
; Sequence 2, Application US/08234783
; Patent No. 5622835
;  GENERAL INFORMATION:
;    APPLICANT:  Herlyn, Meenhard
;    APPLICANT:  Morris, Jennifer
;    APPLICANT:  Rauscher III, Frank J.
;    APPLICANT:  Rodeck, Ulrich
;    TITLE OF INVENTION:  WT1 Monoclonal Antibodies and Methods of
;    TITLE OF INVENTION:  Use Therefor
;    NUMBER OF SEQUENCES:  7
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Howson and Howson
;      STREET:  Spring House Corporate Cntr, PO Box 457
;      CITY:  Spring House
;      STATE:  Pennsylvania
;      COUNTRY:  USA
;      ZIP:  19477
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/234,783
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Bak, Mary E.
;      REGISTRATION NUMBER:  31,215
;      REFERENCE/DOCKET NUMBER:  WST48USA
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  215-540-9200
;      TELEFAX:  215-540-5818
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  210 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-234-783-2


  Query Match             100.0%;  Score 51;  DB 1;  Length 210;
  Best Local Similarity   100.0%;  Pred. No. 0.042;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 RMFPNAPYL 9
          |||||||||
Db    137 RMFPNAPYL 145


RESULT   3
US-08-456-907-2
; Sequence 2, Application US/08456907
; Patent No. 5633142
;  GENERAL INFORMATION:
;    APPLICANT:  Herlyn, Meenhard
;    APPLICANT:  Morris, Jennifer
;    APPLICANT:  Rauscher III, Frank J.
;    APPLICANT:  Rodeck, Ulrich
;    TITLE OF INVENTION:  WT1 Monoclonal Antibodies and Methods of
```

```
;    TITLE OF INVENTION:  Use Therefor
;    NUMBER OF SEQUENCES:  7
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Howson and Howson
;      STREET:  Spring House Corporate Cntr, PO Box 457
;      CITY:  Spring House
;      STATE:  Pennsylvania
;      COUNTRY:  USA
;      ZIP:  19477
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/456,907
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/234,783
;      FILING DATE:  28-APR-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Bak, Mary E.
;      REGISTRATION NUMBER:  31,215
;      REFERENCE/DOCKET NUMBER:  WST48AUSA
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  215-540-9200
;      TELEFAX:  215-540-5818
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  210 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-456-907-2


  Query Match             100.0%;  Score 51;  DB 1;  Length 210;
  Best Local Similarity   100.0%;  Pred. No. 0.042;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 RMFPNAPYL 9
          |||||||||
Db    137 RMFPNAPYL 145


RESULT   4
PCT-US95-05523-2
; Sequence 2, Application PC/TUS9505523
;  GENERAL INFORMATION:
;    APPLICANT:  The Wistar Institute of Anatomy and Biology
;    TITLE OF INVENTION:  WT1 Monoclonal Antibodies and
;    TITLE OF INVENTION:  Methods of Use Therefor
;    NUMBER OF SEQUENCES:  7
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Howson and Howson
;      STREET:  Spring House Corporate Cntr, PO Box 457
;      CITY:  Spring House
;      STATE:  Pennsylvania
;      COUNTRY:  USA
;      ZIP:  19477
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/05523
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/234,783
;      FILING DATE:  28-APR-1994
```

```
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Bak, Mary E.
;      REGISTRATION NUMBER:  31,215
;      REFERENCE/DOCKET NUMBER:  WST48PCT
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  215-540-9200
;      TELEFAX:  215-540-5818
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  210 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
PCT-US95-05523-2


  Query Match             100.0%;  Score 51;  DB 5;  Length 210;
  Best Local Similarity   100.0%;  Pred. No. 0.042;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 RMFPNAPYL 9
          |||||||||
Db    137 RMFPNAPYL 145


RESULT   5
US-08-102-942A-2
; Sequence 2, Application US/08102942A
; Patent No. 5726288
;  GENERAL INFORMATION:
;    APPLICANT:  Call, Katherine M.
;    APPLICANT:  Glaser, Thomas M.
;    APPLICANT:  Ito, Caryn Y.
;    APPLICANT:  Buckler, Alan J.
;    APPLICANT:  Pelletier, Jerry
;    APPLICANT:  Haber, Daniel A.
;    APPLICANT:  Rose, Elise A.
;    APPLICANT:  Housman, David E.
;    APPLICANT:  Bruening, Wendy
;    APPLICANT:  Darveau, Andre
;    TITLE OF INVENTION:  Localization and Characterization of the
;    TITLE OF INVENTION:  Wilms' Tumor Gene
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Hamilton, Brook, Smith & Reynolds, P.C.
;      STREET:  Two Militia Drive
;      CITY:  Lexington
;      STATE:  Massachusetts
;      COUNTRY:  U.S.
;      ZIP:  02173
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/102,942A
;      FILING DATE:  02-AUG-1993
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Granahan, Patricia
;      REGISTRATION NUMBER:  32,227
;      REFERENCE/DOCKET NUMBER:  MIT-5194A2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-861-6240
;      TELEFAX:  617-861-9540
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  345 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-102-942A-2
```

```
  Query Match             100.0%;  Score 51;  DB 1;  Length 345;
  Best Local Similarity   100.0%;  Pred. No. 0.071;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 RMFPNAPYL 9
          |||||||||
Db     42 RMFPNAPYL 50


RESULT   6
US-08-234-783-4
; Sequence 4, Application US/08234783
; Patent No. 5622835
;  GENERAL INFORMATION:
;    APPLICANT:  Herlyn, Meenhard
;    APPLICANT:  Morris, Jennifer
;    APPLICANT:  Rauscher III, Frank J.
;    APPLICANT:  Rodeck, Ulrich
;    TITLE OF INVENTION:  WT1 Monoclonal Antibodies and Methods of
;    TITLE OF INVENTION:  Use Therefor
;    NUMBER OF SEQUENCES:  7
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Howson and Howson
;      STREET:  Spring House Corporate Cntr, PO Box 457
;      CITY:  Spring House
;      STATE:  Pennsylvania
;      COUNTRY:  USA
;      ZIP:  19477
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/234,783
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Bak, Mary E.
;      REGISTRATION NUMBER:  31,215
;      REFERENCE/DOCKET NUMBER:  WST48USA
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  215-540-9200
;      TELEFAX:  215-540-5818
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  429 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-234-783-4


  Query Match             100.0%;  Score 51;  DB 1;  Length 429;
  Best Local Similarity   100.0%;  Pred. No. 0.09;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 RMFPNAPYL 9
          |||||||||
Db    126 RMFPNAPYL 134


RESULT   7
US-08-456-907-4
; Sequence 4, Application US/08456907
; Patent No. 5633142
;  GENERAL INFORMATION:
;    APPLICANT:  Herlyn, Meenhard
;    APPLICANT:  Morris, Jennifer
;    APPLICANT:  Rauscher III, Frank J.
;    APPLICANT:  Rodeck, Ulrich
```

```
;    TITLE OF INVENTION:  WT1 Monoclonal Antibodies and Methods of
;    TITLE OF INVENTION:  Use Therefor
;    NUMBER OF SEQUENCES:  7
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Howson and Howson
;      STREET:  Spring House Corporate Cntr, PO Box 457
;      CITY:  Spring House
;      STATE:  Pennsylvania
;      COUNTRY:  USA
;      ZIP:  19477
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/456,907
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/234,783
;      FILING DATE:  28-APR-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Bak, Mary E.
;      REGISTRATION NUMBER:  31,215
;      REFERENCE/DOCKET NUMBER:  WST48AUSA
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  215-540-9200
;      TELEFAX:  215-540-5818
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  429 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-456-907-4


  Query Match             100.0%;  Score 51;  DB 1;  Length 429;
  Best Local Similarity   100.0%;  Pred. No. 0.09;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 RMFPNAPYL 9
          |||||||||
Db    126 RMFPNAPYL 134


RESULT   8
PCT-US95-05523-4
; Sequence 4, Application PC/TUS9505523
;  GENERAL INFORMATION:
;    APPLICANT:  The Wistar Institute of Anatomy and Biology
;    TITLE OF INVENTION:  WT1 Monoclonal Antibodies and
;    TITLE OF INVENTION:  Methods of Use Therefor
;    NUMBER OF SEQUENCES:  7
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Howson and Howson
;      STREET:  Spring House Corporate Cntr, PO Box 457
;      CITY:  Spring House
;      STATE:  Pennsylvania
;      COUNTRY:  USA
;      ZIP:  19477
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/05523
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/234,783
;      FILING DATE:  28-APR-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Bak, Mary E.
;      REGISTRATION NUMBER:  31,215
;      REFERENCE/DOCKET NUMBER:  WST48PCT
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  215-540-9200
;      TELEFAX:  215-540-5818
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  429 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
PCT-US95-05523-4


  Query Match             100.0%;  Score 51;  DB 5;  Length 429;
  Best Local Similarity   100.0%;  Pred. No. 0.09;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 RMFPNAPYL 9
          |||||||||
Db    126 RMFPNAPYL 134


RESULT   9
US-08-102-942A-4
; Sequence 4, Application US/08102942A
; Patent No. 5726288
;  GENERAL INFORMATION:
;    APPLICANT:  Call, Katherine M.
;    APPLICANT:  Glaser, Thomas M.
;    APPLICANT:  Ito, Caryn Y.
;    APPLICANT:  Buckler, Alan J.
;    APPLICANT:  Pelletier, Jerry
;    APPLICANT:  Haber, Daniel A.
;    APPLICANT:  Rose, Elise A.
;    APPLICANT:  Housman, David E.
;    APPLICANT:  Bruening, Wendy
;    APPLICANT:  Darveau, Andre
;    TITLE OF INVENTION:  Localization and Characterization of the
;    TITLE OF INVENTION:  Wilms' Tumor Gene
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Hamilton, Brook, Smith & Reynolds, P.C.
;      STREET:  Two Militia Drive
;      CITY:  Lexington
;      STATE:  Massachusetts
;      COUNTRY:  U.S.
;      ZIP:  02173
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/102,942A
;      FILING DATE:  02-AUG-1993
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Granahan, Patricia
;      REGISTRATION NUMBER:  32,227
;      REFERENCE/DOCKET NUMBER:  MIT-5194A2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-861-6240
;      TELEFAX:  617-861-9540
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  449 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
```

```
US-08-102-942A-4

  Query Match             100.0%;  Score 51;  DB 1;  Length 449;
  Best Local Similarity   100.0%;  Pred. No. 0.094;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RMFPNAPYL 9
           |||||||||
Db     126 RMFPNAPYL 134


RESULT  10
US-08-102-942A-6
; Sequence 6, Application US/08102942A
; Patent No. 5726288
;  GENERAL INFORMATION:
;    APPLICANT:  Call, Katherine M.
;    APPLICANT:  Glaser, Thomas M.
;    APPLICANT:  Ito, Caryn Y.
;    APPLICANT:  Buckler, Alan J.
;    APPLICANT:  Pelletier, Jerry
;    APPLICANT:  Haber, Daniel A.
;    APPLICANT:  Rose, Elise A.
;    APPLICANT:  Housman, David E.
;    APPLICANT:  Bruening, Wendy
;    APPLICANT:  Darveau, Andre
;    TITLE OF INVENTION:  Localization and Characterization of the
;    TITLE OF INVENTION:  Wilms' Tumor Gene
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Hamilton, Brook, Smith & Reynolds, P.C.
;      STREET:  Two Militia Drive
;      CITY:  Lexington
;      STATE:  Massachusetts
;      COUNTRY:  U.S.
;      ZIP:  02173
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/102,942A
;      FILING DATE:  02-AUG-1993
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Granahan, Patricia
;      REGISTRATION NUMBER:  32,227
;      REFERENCE/DOCKET NUMBER:  MIT-5194A2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-861-6240
;      TELEFAX:  617-861-9540
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  449 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  not relevant
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-102-942A-6


  Query Match             100.0%;  Score 51;  DB 1;  Length 449;
  Best Local Similarity   100.0%;  Pred. No. 0.094;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RMFPNAPYL 9
           |||||||||
Db     126 RMFPNAPYL 134


RESULT  11
US-08-975-080-26
; Sequence 26, Application US/08975080
; Patent No. 6245523
;  GENERAL INFORMATION:
;    APPLICANT:  Altieri, Dario C.
;    TITLE OF INVENTION:  SURVIVIN, A PROTEIN THAT INHIBITS
;    TITLE OF INVENTION:  CELLULAR APOPTOSIS, AND ITS MODULATION
;    NUMBER OF SEQUENCES:  35
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MORGAN, LEWIS & BOCKIUS LLP
;      STREET:  1800 M Street, N.W.
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20036-5869
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/975,080
;      FILING DATE:  20-NOV-1997
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/031,435
;      FILING DATE:  20-NOV-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Adler, Reid G.
;      REGISTRATION NUMBER:  30,988
;      REFERENCE/DOCKET NUMBER:  044574-5022-01-WO
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-467-7000
;      TELEFAX:  202-467-7176
;  INFORMATION FOR SEQ ID NO:  26:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  50 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-975-080-26


  Query Match              72.5%;  Score 37;  DB 4;  Length 50;
  Best Local Similarity    66.7%;  Pred. No. 2.8;
  Matches    6;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 RMFPNAPYL 9
           | ||| |:|
Db      32 RCFPNCPFL 40


RESULT  12
US-08-836-134-23
; Sequence 23, Application US/08836134A
; Patent No. 6020127
; GENERAL INFORMATION:
;  APPLICANT: Mackenzie, Alex E.
;  APPLICANT: Korneluk, Robert G.
;  APPLICANT: Mahadevan, Mani S.
;  APPLICANT: McLean, Michael
;  APPLICANT: Roy, Natalie
;  APPLICANT: Ikeda, Joh-e
;  TITLE OF INVENTION: Neuronal Apoptosis Inhibitor Protein, Gene Sequence and
; Patent No. 6020127
;  TITLE OF INVENTION: Mutations Causative of Spinal Muscular Atrophy
;  FILE REFERENCE: 3477-112, 033477/139914
;  CURRENT APPLICATION NUMBER: US/08/836,134A
;  CURRENT FILING DATE: 1997-06-20
;  NUMBER OF SEQ ID NOS: 23
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 23
;   LENGTH: 1151
```

```
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-08-836-134-23


  Query Match             72.5%;  Score 37;  DB 3;  Length 1151;
  Best Local Similarity   66.7%;  Pred. No. 76;
  Matches    6;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 RMFPNAPYL 9
           | ||| |:|
Db     346 RCFPNCPFL 354


RESULT 13
US-08-836-134-2
; Sequence 2, Application US/08836134A
; Patent No. 6020127
; GENERAL INFORMATION:
;  APPLICANT: Mackenzie, Alex E.
;  APPLICANT: Korneluk, Robert G.
;  APPLICANT: Mahadevan, Mani S.
;  APPLICANT: McLean, Michael
;  APPLICANT: Roy, Natalie
;  APPLICANT: Ikeda, Joh-e
;  TITLE OF INVENTION: Neuronal Apoptosis Inhibitor Protein, Gene Sequence and
; Patent No. 6020127
;  TITLE OF INVENTION: Mutations Causative of Spinal Muscular Atrophy
;  FILE REFERENCE: 3477-112, 033477/139914
;  CURRENT APPLICATION NUMBER: US/08/836,134A
;  CURRENT FILING DATE: 1997-06-20
;  NUMBER OF SEQ ID NOS: 23
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
;   LENGTH: 1232
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-08-836-134-2


  Query Match             72.5%;  Score 37;  DB 3;  Length 1232;
  Best Local Similarity   66.7%;  Pred. No. 81;
  Matches    6;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 RMFPNAPYL 9
           | ||| |:|
Db     337 RCFPNCPFL 345


RESULT 14
US-08-511-485-23
; Sequence 23, Application US/08511485
; Patent No. 5919912
;  GENERAL INFORMATION:
;    APPLICANT:  Korneluk, Robert G.
;    APPLICANT:  Mackenzie, Alexander E.
;    APPLICANT:  Baird, Stephen
;    TITLE OF INVENTION:  MAMMALIAN IAP GENE FAMILY, PRIMERS,
;    TITLE OF INVENTION:  PROBES, AND DETECTION METHODS
;    NUMBER OF SEQUENCES:  38
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson P.C.
;      STREET:  225 Franklin Street
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  02110-2804
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
```

```
;      APPLICATION NUMBER:  US/08/511,485
;      FILING DATE:  04-AUG-1995
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Clark, Paul T.
;      REGISTRATION NUMBER:  30,162
;      REFERENCE/DOCKET NUMBER:  07540/002001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617/542-5070
;      TELEFAX:  617/542-8906
;      TELEX:  200154
;  INFORMATION FOR SEQ ID NO:  23:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  67 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  not relevant
;      TOPOLOGY:  both
;    MOLECULE TYPE:  protein
US-08-511-485-23


  Query Match             70.6%;  Score 36;  DB 2;  Length 67;
  Best Local Similarity   66.7%;  Pred. No. 5.8;
  Matches    6;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 RMFPNAPYL 9
           | ||| |:|
Db      59 RHFPNCPFL 67


RESULT 15
US-08-297-431B-33
; Sequence 33, Application US/08297431B
; Patent No. 6136605
;  GENERAL INFORMATION:
;    APPLICANT:  Fahl, William E
;    APPLICANT:  Gulick, Andrew M
;    APPLICANT:  Manoharan, T Herbert
;    APPLICANT:  Puchalski, Ralph B
;    APPLICANT:  Kramer, Katharine
;    APPLICANT:  Wasserman Wyeth W
;    TITLE OF INVENTION:  NOVEL GLUTATHIONE S-TRANSFERASE ISOFORMS
;    NUMBER OF SEQUENCES:  34
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Saul, Ewing, Remick & Saul, LLP
;      STREET:  1500 Market Street, 38th Floor
;      CITY:  Philadelphia
;      STATE:  PA
;      COUNTRY:  US
;      ZIP:  19102-2186
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/297,431B
;      FILING DATE:  August 26, 1994
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Reed, Janet E.
;      REGISTRATION NUMBER:  36252
;      REFERENCE/DOCKET NUMBER:  WARF F039
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  215-972-8386
;      TELEFAX:  215-972-2292
;  INFORMATION FOR SEQ ID NO:  33:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  217 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-297-431B-33
```

```
  Query Match             70.6%;  Score 36;  DB 4;  Length 217;
  Best Local Similarity   85.7%;  Pred. No. 20;
  Matches    6;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       3 FPNAPYL 9
           ||| |||
Db      56 FPNLPYL 62


Search completed: October 23, 2001, 13:21:32
Job time: 85 sec
```

```
                      GenCore version 4.5
            Copyright (c) 1993 - 2000  Compugen Ltd.


OM protein - protein search, using sw model

Run on:          October 23, 2001, 13:22:52 ; Search time 18.81 Seconds
                                              (without alignments)
                                              9.852 Million cell updates/sec

Title:           US-09-625-963-1
Perfect score:   51
Sequence:        1 RMFPNAPYL 9

Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched:        197339 seqs, 20590346 residues

Total number of hits satisfying chosen parameters:      50818

Minimum DB seq length: 0
Maximum DB seq length: 9

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       Issued_Patents_AA:*
                1:  /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
                2:  /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
                3:  /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
                4:  /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
                5:  /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
                6:  /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                              SUMMARIES
                   %
Result           Query
  No.    Score   Match Length DB  ID                     Description
  ---------------------------------------------------------------------
    1      28     54.9      9  3  US-09-162-368B-26       Sequence 26, Appl
    2      28     54.9      9  4  US-09-161-877B-26       Sequence 26, Appl
    3      27     52.9      8  2  US-08-737-085A-6        Sequence 6, Appli
    4      27     52.9      8  3  US-09-246-258-6         Sequence 6, Appli
    5      27     52.9      8  4  US-09-532-106-6         Sequence 6, Appli
    6      26     51.0      6  2  US-08-672-805-12        Sequence 12, Appl
    7      25     49.0      8  5  PCT-US94-03744-6        Sequence 6, Appli
    8      25     49.0      9  3  US-09-162-368B-28       Sequence 28, Appl
    9      25     49.0      9  4  US-09-161-877B-28       Sequence 28, Appl
   10      24     47.1      7  1  US-08-261-525A-6        Sequence 6, Appli
   11      24     47.1      9  2  US-08-417-174-13        Sequence 13, Appl
   12      24     47.1      9  2  US-08-231-565A-13       Sequence 13, Appl
   13      24     47.1      9  2  US-09-007-961-13        Sequence 13, Appl
   14      24     47.1      9  3  US-08-159-339A-393      Sequence 393, App
   15      24     47.1      9  3  US-09-162-368B-24       Sequence 24, Appl
   16      24     47.1      9  4  US-09-161-877B-24       Sequence 24, Appl
   17      24     47.1      9  4  US-08-786-455B-5        Sequence 5, Appli
   18      23     45.1      7  6  5514590-12             Patent No. 5514590
   19      23     45.1      8  2  US-08-177-109A-24       Sequence 24, Appl
   20      23     45.1      8  2  US-08-687-706-24        Sequence 24, Appl
   21      23     45.1      8  4  US-08-444-818-433       Sequence 433, App
   22      23     45.1      8  4  US-08-444-818-434       Sequence 434, App
   23      23     45.1      8  5  PCT-US94-01321-70       Sequence 70, Appl
   24      22     43.1      6  2  US-08-672-805-11        Sequence 11, Appl
   25      22     43.1      8  4  US-09-484-318-7         Sequence 7, Appli
   26      22     43.1      8  4  US-09-484-319-7         Sequence 7, Appli
   27      22     43.1      8  4  US-09-484-320-7         Sequence 7, Appli
```

```
   28      22     43.1      8  4  US-09-484-321-7         Sequence 7, Appli
   29      22     43.1      8  4  US-09-484-323-7         Sequence 7, Appli
   30      22     43.1      8  4  US-09-325-769-8         Sequence 8, Appli
   31      22     43.1      9  3  US-09-162-368B-25       Sequence 25, Appl
   32      22     43.1      9  4  US-09-161-877B-25       Sequence 25, Appl
   33      21     41.2      5  1  US-07-880-216-1         Sequence 1, Appli
   34      21     41.2      7  2  US-08-968-676-158       Sequence 158, App
   35      21     41.2      7  5  PCT-US94-01321-69       Sequence 69, Appl
   36      21     41.2      8  1  US-08-594-447-37        Sequence 37, Appl
   37      21     41.2      8  1  US-08-541-964-36        Sequence 36, Appl
   38      21     41.2      8  2  US-08-665-647-51        Sequence 51, Appl
   39      21     41.2      9  2  US-08-725-736D-12       Sequence 12, Appl
   40      21     41.2      9  3  US-09-162-368B-12       Sequence 12, Appl
   41      21     41.2      9  3  US-09-162-368B-30       Sequence 30, Appl
   42      21     41.2      9  3  US-08-599-226-14        Sequence 14, Appl
   43      21     41.2      9  3  US-08-599-226-19        Sequence 19, Appl
   44      21     41.2      9  4  US-09-161-877B-12       Sequence 12, Appl
   45      21     41.2      9  4  US-09-161-877B-30       Sequence 30, Appl


                              ALIGNMENTS


RESULT   1
US-09-162-368B-26
; Sequence 26, Application US/09162368B
; Patent No. 6083703
;  GENERAL INFORMATION:
;    APPLICANT:  WANG, R.F.; ROSENBERG, S. A.
;    TITLE OF INVENTION:  IDENTIFICATION OF TRP-2 AS
;    TITLE OF INVENTION:  A HUMAN TUMOR ANTIGEN RECOGNIZED BY CYTOTOXIC T
;    TITLE OF INVENTION:  LYMPHOCYTES
;    NUMBER OF SEQUENCES:  31
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MORGAN & FINNEGAN, L.L.P.
;      STREET:  345 PARK AVENUE
;      CITY:  NEW YORK
;      STATE:  NEW YORK
;      COUNTRY:  USA
;      ZIP:  10154
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  FLOPPY DISK
;      COMPUTER:  IBM PC COMPATIBLE
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  MICROSOFT WORD 97
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/162,368B
;      FILING DATE:  28-SEPT-1998
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/725,736
;      FILING DATE:  04-OCT-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/599,602
;      FILING DATE:  09-FEB-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  KATHRYN M. BROWN
;      REGISTRATION NUMBER:  34,556
;      REFERENCE/DOCKET NUMBER:  2026-4243US1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 758-4800
;      TELEFAX:  (212) 751-6849
;      TELEX:  421792
;  INFORMATION FOR SEQ ID NO:  26:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9
;      TYPE:  AMINO ACID
;      STRANDEDNESS:  UNKNOWN
;      TOPOLOGY:  UNKNOWN
;    MOLECULE TYPE:
;      DESCRIPTION:  PEPTIDE
;    FEATURE:
;      NAME/KEY:
```

```
;      LOCATION:
;      IDENTIFICATION METHOD:
;      OTHER INFORMATION:
US-09-162-368B-26

  Query Match             54.9%;  Score 28;  DB 3;  Length 9;
  Best Local Similarity   57.1%;  Pred. No. 1.5e+05;
  Matches    4;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      2 MFPNAPY 8
          :||  ||
Db      2 LFPGRPY 8

RESULT   2
US-09-161-877B-26
; Sequence 26, Application US/09161877B
; Patent No. 6132980
;  GENERAL INFORMATION:
;    APPLICANT:  WANG, R.F.; ROSENBERG, S. A.
;    TITLE OF INVENTION:  IDENTIFICATION OF TRP-2
;    TITLE OF INVENTION:  AS A HUMAN TUMOR ANTIGEN RECOGNIZED BY
;    TITLE OF INVENTION:  CYTOTOXIC T LYMPHOCYTES
;    NUMBER OF SEQUENCES:  32
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MORGAN & FINNEGAN, L.L.P.
;      STREET:  345 PARK AVENUE
;      CITY:  NEW YORK
;      STATE:  NEW YORK
;      COUNTRY:  USA
;      ZIP:  10154
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  FLOPPY DISK
;      COMPUTER:  IBM PC COMPATIBLE
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  MICROSOFT WORD 97
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/161,877B
;      FILING DATE:  28-SEPT-1998
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/725,736
;      FILING DATE:  04-OCT-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/599,602
;      FILING DATE:  09-FEB-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  KATHRYN M. BROWN
;      REGISTRATION NUMBER:  34,556
;      REFERENCE/DOCKET NUMBER:  2026-4243US2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 758-4800
;      TELEFAX:  (212) 751-6849
;      TELEX:  421792
;  INFORMATION FOR SEQ ID NO:  26:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9
;      TYPE:  AMINO ACID
;      STRANDEDNESS:  UNKNOWN
;      TOPOLOGY:  UNKNOWN
;    MOLECULE TYPE:
;      DESCRIPTION:  PEPTIDE
;    FEATURE:
;      NAME/KEY:
;      LOCATION:
;      IDENTIFICATION METHOD:
;      OTHER INFORMATION:
US-09-161-877B-26

  Query Match             54.9%;  Score 28;  DB 4;  Length 9;
  Best Local Similarity   57.1%;  Pred. No. 1.5e+05;
  Matches    4;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      2 MFPNAPY 8
          :||  ||
Db      2 LFPGRPY 8

RESULT   3
US-08-737-085A-6
; Sequence 6, Application US/08737085A
; Patent No. 5869232
;  GENERAL INFORMATION:
;    APPLICANT:  SALLBERG, MATTI
;    TITLE OF INVENTION:  ANTIGEN/ANTIBODY SPECIFICITY
;    TITLE OF INVENTION:  EXCHANGER
;    NUMBER OF SEQUENCES:  23
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  DARBY & DARBY PC
;      STREET:  805 Third Avenue
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10022
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/737,085A
;      FILING DATE:  27-DEC-1996
;      CLASSIFICATION:  426
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Green, Reza
;      REGISTRATION NUMBER:  38,475
;      REFERENCE/DOCKET NUMBER:  3846/0C569
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-527-7659
;      TELEFAX:  212-753-6237
;      TELEX:  236687
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  8 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-737-085A-6

  Query Match             52.9%;  Score 27;  DB 2;  Length 8;
  Best Local Similarity   83.3%;  Pred. No. 1.5e+05;
  Matches    5;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      4 PNAPYL 9
          |||| |
Db      2 PNAPIL 7

RESULT   4
US-09-246-258-6
; Sequence 6, Application US/09246258
; Patent No. 6040137
;  GENERAL INFORMATION:
;    APPLICANT:  SALLBERG, MATTI
;    TITLE OF INVENTION:  ANTIGEN/ANTIBODY SPECIFICITY
;    TITLE OF INVENTION:  EXCHANGER
;    NUMBER OF SEQUENCES:  23
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  DARBY & DARBY PC
```

```
;      STREET:  805 Third Avenue
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10022
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/246,258
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/737,085
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Green, Reza
;      REGISTRATION NUMBER:  38,475
;      REFERENCE/DOCKET NUMBER:  3846/0C569
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-527-7659
;      TELEFAX:  212-753-6237
;      TELEX:  236687
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  8 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-09-246-258-6

  Query Match             52.9%;  Score 27;  DB 3;  Length 8;
  Best Local Similarity   83.3%;  Pred. No. 1.5e+05;
  Matches    5;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      4 PNAPYL 9
          |||| |
Db      2 PNAPIL 7


RESULT   5
US-09-532-106-6
; Sequence 6, Application US/09532106
; Patent No. 6245895
;   GENERAL INFORMATION:
;        APPLICANT: SALLBERG, MATTI
;        TITLE OF INVENTION: ANTIGEN/ANTIBODY SPECIFICITY
;                            EXCHANGER
;        NUMBER OF SEQUENCES: 23
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: DARBY & DARBY PC
;             STREET: 805 Third Avenue
;             CITY: New York
;             STATE: New York
;             COUNTRY: USA
;             ZIP: 10022
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Diskette
;             COMPUTER: IBM Compatible
;             OPERATING SYSTEM: DOS
;             SOFTWARE: FastSEQ for Windows Version 2.0
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/532,106
;             FILING DATE: 21-Mar-2000
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US/08/737,085A
;             FILING DATE: 27-DEC-1996
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Green, Reza
```

```
;             REGISTRATION NUMBER: 38,475
;             REFERENCE/DOCKET NUMBER: 3846/0C569
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 212-527-7659
;             TELEFAX: 212-753-6237
;             TELEX: 236687
;     INFORMATION FOR SEQ ID NO: 6:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 8 amino acids
;             TYPE: amino acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: peptide
;        SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-532-106-6

  Query Match             52.9%;  Score 27;  DB 4;  Length 8;
  Best Local Similarity   83.3%;  Pred. No. 1.5e+05;
  Matches    5;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      4 PNAPYL 9
          |||| |
Db      2 PNAPIL 7


RESULT   6
US-08-672-805-12
; Sequence 12, Application US/08672805
; Patent No. 5831003
;  GENERAL INFORMATION:
;    APPLICANT:  Baumbach, George A.,
;    APPLICANT:  Buettner, Joseph A.,
;    APPLICANT:  Dadd, Christopher A.,
;    APPLICANT:  Hammond, David J.
;    TITLE OF INVENTION:  Peptides Which Bind to Prothrombin and
;    TITLE OF INVENTION:  Thrombin
;    NUMBER OF SEQUENCES:  21
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Bayer Corporation
;      STREET:  800 Dwight Way
;      STREET:  P. O. Box 1986
;      CITY:  Berkeley
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94701-1986
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.50 inch, 1.44Mb Storage
;      COMPUTER:  IBM
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  WordPerfect 6.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/672,805
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Giblin, James A.
;      REGISTRATION NUMBER:  25772
;      REFERENCE/DOCKET NUMBER:  MSB-7236
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (510)705-7910
;      TELEFAX:  (510)705-7904
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  6
;      TYPE:  amino acid
;      STRANDEDNESS:  single strand
;      TOPOLOGY:  linear
;    MOLECULE TYPE:
;      DESCRIPTION:  peptide
```

```
US-08-672-805-12

  Query Match             51.0%;  Score 26;  DB 2;  Length 6;
  Best Local Similarity   83.3%;  Pred. No. 1.5e+05;
  Matches    5;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       4 PNAPYL 9
           | ||||
Db       1 PFAPYL 6


RESULT   7
PCT-US94-03744-6
; Sequence 6, Application PC/TUS9403744
;  GENERAL INFORMATION:
;    APPLICANT:  PETRI, WILLIAM A.
;    APPLICANT:  McCOY, JAMES J.
;    APPLICANT:  MANN, BARBARA J.
;    TITLE OF INVENTION:  35/31 KDA SUBUNIT OF THE
;    TITLE OF INVENTION:  ENTAMOEBA
;    TITLE OF INVENTION:  HISTOLYTICA ADHERENCE LECTIN
;    NUMBER OF SEQUENCES:  11
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Morrison & Foerster
;      STREET:  2000 Pennsylvania Avenue, Suite 5500
;      CITY:  Washington
;      STATE:  DC
;      COUNTRY:  USA
;      ZIP:  20006-1812
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US94/03744
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/045,679
;      FILING DATE:  09-APR-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  MURASHIGE, KATE H.
;      REGISTRATION NUMBER:  29,959
;      REFERENCE/DOCKET NUMBER:  29148-20005.00
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 887-1500
;      TELEFAX:  (202) 887-0763
;      TELEX:  90-4030
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  8 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
PCT-US94-03744-6


  Query Match             49.0%;  Score 25;  DB 5;  Length 8;
  Best Local Similarity   80.0%;  Pred. No. 1.5e+05;
  Matches    4;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       4 PNAPY 8
             || ||
Db       3 PNYPY 7


RESULT   8
US-09-162-368B-28
; Sequence 28, Application US/09162368B
; Patent No. 6003703
;  GENERAL INFORMATION:
```

```
;    APPLICANT:  WANG, R.F.; ROSENBERG, S. A.
;    TITLE OF INVENTION:  IDENTIFICATION OF TRP-2 AS
;    TITLE OF INVENTION:  A HUMAN TUMOR ANTIGEN RECOGNIZED BY CYTOTOXIC T
;    TITLE OF INVENTION:  LYMPHOCYTES
;    NUMBER OF SEQUENCES:  31
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MORGAN & FINNEGAN, L.L.P.
;      STREET:  345 PARK AVENUE
;      CITY:  NEW YORK
;      STATE:  NEW YORK
;      COUNTRY:  USA
;      ZIP:  10154
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  FLOPPY DISK
;      COMPUTER:  IBM PC COMPATIBLE
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  MICROSOFT WORD 97
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/162,368B
;      FILING DATE:  28-SEPT-1998
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/725,736
;      FILING DATE:  04-OCT-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/599,602
;      FILING DATE:  09-FEB-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  KATHRYN M. BROWN
;      REGISTRATION NUMBER:  34,556
;      REFERENCE/DOCKET NUMBER:  2026-4243US1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 758-4800
;      TELEFAX:  (212) 751-6849
;      TELEX:  421792
;  INFORMATION FOR SEQ ID NO:  28:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9
;      TYPE:  AMINO ACID
;      STRANDEDNESS:  UNKNOWN
;      TOPOLOGY:  UNKNOWN
;    MOLECULE TYPE:
;      DESCRIPTION:  PEPTIDE
;    FEATURE:
;      NAME/KEY:
;      LOCATION:
;      IDENTIFICATION METHOD:
;      OTHER INFORMATION:
US-09-162-368B-28


  Query Match             49.0%;  Score 25;  DB 3;  Length 9;
  Best Local Similarity   80.0%;  Pred. No. 1.5e+05;
  Matches    4;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       4 PNAPY 8
           | |||
Db       4 PGAPY 8


RESULT   9
US-09-161-877B-28
; Sequence 28, Application US/09161877B
; Patent No. 6132980
;  GENERAL INFORMATION:
;    APPLICANT:  WANG, R.F.; ROSENBERG, S. A.
;    TITLE OF INVENTION:  IDENTIFICATION OF TRP-2
;    TITLE OF INVENTION:  AS A HUMAN TUMOR ANTIGEN RECOGNIZED BY
;    TITLE OF INVENTION:  CYTOTOXIC T LYMPHOCYTES
;    NUMBER OF SEQUENCES:  32
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MORGAN & FINNEGAN, L.L.P.
;      STREET:  345 PARK AVENUE
```

```
;      CITY:  NEW YORK
;      STATE:  NEW YORK
;      COUNTRY:  USA
;      ZIP:  10154
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  FLOPPY DISK
;      COMPUTER:  IBM PC COMPATIBLE
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  MICROSOFT WORD 97
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/161,877B
;      FILING DATE:  28-SEPT-1998
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/725,736
;      FILING DATE:  04-OCT-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/599,602
;      FILING DATE:  09-FEB-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  KATHRYN M. BROWN
;      REGISTRATION NUMBER:  34,556
;      REFERENCE/DOCKET NUMBER:  2026-4243US2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 758-4800
;      TELEFAX:  (212) 751-6849
;      TELEX:  421792
;  INFORMATION FOR SEQ ID NO:  28:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9
;      TYPE:  AMINO ACID
;      STRANDEDNESS:  UNKNOWN
;      TOPOLOGY:  UNKNOWN
;    MOLECULE TYPE:
;      DESCRIPTION:  PEPTIDE
;    FEATURE:
;      NAME/KEY:
;      LOCATION:
;      IDENTIFICATION METHOD:
;      OTHER INFORMATION:
US-09-161-877B-28


  Query Match             49.0%;  Score 25;  DB 4;  Length 9;
  Best Local Similarity   80.0%;  Pred. No. 1.5e+05;
  Matches    4;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      4 PNAPY 8
          | |||
Db      4 PGAPY 8


RESULT  10
US-08-261-525A-6
; Sequence 6, Application US/08261525A
; Patent No. 5569598
;  GENERAL INFORMATION:
;    APPLICANT:  PARK, Soon Jae
;    APPLICANT:  LEE, Young Mee
;    APPLICANT:  WON, Teug Yeon
;    APPLICANT:  KWON, Soon Chang
;    APPLICANT:  LEE, Seung Joo
;    APPLICANT:  KIM, Jung Ho
;    APPLICANT:  KIM, Bum Joon
;    TITLE OF INVENTION:  NOVEL AMINOPEPTIDASE, PROCESSES FOR
;    TITLE OF INVENTION:  THE PREPARATION OF AMINOPEPTIDASE AND
;    TITLE OF INVENTION:  PROTEIN THEREFROM
;    NUMBER OF SEQUENCES:  6
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  PARK, Soon Jae
;      STREET:  Lucky Apt. 6-101, Doryong-dong, Yuseong-gu
;      CITY:  Daejeon
;      STATE:  Daejeon
```

```
;      COUNTRY:  Republic of Korea
;      ZIP:  305-340
;      ADDRESSEE:  LEE, Young Mee
;      STREET:  Shinsung Apt. 3-306, 28-4, Nae-dong, Seo-gu
;      CITY:  Daejeon
;      STATE:  Daejeon
;      COUNTRY:  Republic of Korea
;      ZIP:  302-181
;      ADDRESSEE:  WON, Teug Yeon
;      STREET:  Gongjak Hanyang Apt. 3-1201, 835,
;      STREET:  Tanbang-dong, Seo-gu
;      CITY:  Daejeon
;      STATE:  Daejeon
;      COUNTRY:  Republic of Korea
;      ZIP:  302-223
;      ADDRESSEE:  KWON, Soon Chang
;      STREET:  Shinsung Apt. 3-107, 28-4, Nae-dong, Seo-gu
;      CITY:  Daejeon
;      STATE:  Daejeon
;      COUNTRY:  Republic of Korea
;      ZIP:  302-181
;      ADDRESSEE:  LEE, Seung Joo
;      STREET:  Lucky Apt. B-107, 386-4, Doryong-dong,
;      STREET:  Yuseong-gu
;      CITY:  Daejeon
;      STATE:  Daejeon
;      COUNTRY:  Republic of Korea
;      ZIP:  305-340
;      ADDRESSEE:  KIM, Jung Ho
;      STREET:  Lucky Dormitory 511, 386-1, Doryong-dong,
;      STREET:  Yuseong-gu
;      CITY:  Daejeon
;      STATE:  Daejeon
;      COUNTRY:  Republic of Korea
;      ZIP:  305-340
;      ADDRESSEE:  KIM, Bum Joon
;      STREET:  Lucky Dormitory 502, 386-1, Doryong-dong,
;      STREET:  Yuseong-gu
;      CITY:  Daejeon
;      STATE:  Daejeon
;      COUNTRY:  Republic of Korea
;      ZIP:  305-340
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk, 3.5 inch, 1.44MB storage
;      COMPUTER:  IBM PC/AT
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  Word Perfect 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/261,525A
;      FILING DATE:  17-JUN-1994
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  KR 93-11107
;      FILING DATE:  17-JUN-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Shahan Islam
;      REGISTRATION NUMBER:  32,507
;      REFERENCE/DOCKET NUMBER:  A-9883
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 278-1000
;      TELEFAX:  (212) 953-7249
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  7 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-261-525A-6


  Query Match             47.1%;  Score 24;  DB 1;  Length 7;
  Best Local Similarity   66.7%;  Pred. No. 1.5e+05;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;
```

```
Qy       2 MFPNAP 7
           |||  |
Db       1 MFPTEP 6


RESULT  11
US-08-417-174-13
; Sequence 13, Application US/08417174
; Patent No. 5844075
;  GENERAL INFORMATION:
;    APPLICANT:  KAWAKAMI, YUTAKA; ROSENBERG,
;    APPLICANT:  STEVEN A.
;    TITLE OF INVENTION:  MELANOMA ANTIGENS AND
;    TITLE OF INVENTION:  THEIR USE IN DIAGNOSTIC AND THERAPEUTIC
;    TITLE OF INVENTION:  METHODS
;    NUMBER OF SEQUENCES:  126
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MORGAN & FINNEGAN, L.L.P.
;      STREET:  345 PARK AVENUE
;      CITY:  NEW YORK
;      STATE:  NEW YORK
;      COUNTRY:  USA
;      ZIP:  10154
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  FLOPPY DISK
;      COMPUTER:  IBM PC COMPATIBLE
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  ASCII
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/417,174
;      FILING DATE:  05-APR-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/231,565
;      FILING DATE:  22-APR-1994
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  CAROL M. GRUPPI
;      REGISTRATION NUMBER:  37,341
;      REFERENCE/DOCKET NUMBER:  2026-4124US1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 758-4800
;      TELEFAX:  (212) 751-6849
;      TELEX:  421792
;  INFORMATION FOR SEQ ID NO:  13:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9
;      TYPE:  amino acid
;      STRANDEDNESS:  Unknown
;      TOPOLOGY:  Unknown
;    MOLECULE TYPE:  Peptide
US-08-417-174-13

  Query Match             47.1%;  Score 24;  DB 2;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 1.5e+05;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       4 PNAP 7
           ||||
Db       4 PNAP 7


RESULT  12
US-08-231-565A-13
; Sequence 13, Application US/08231565A
; Patent No. 5874560
;  GENERAL INFORMATION:
;    APPLICANT:  KAWAKAMI, YUTAKA; ROSENBERG,
;    APPLICANT:  STEVEN A.
;    TITLE OF INVENTION:  MELANOMA ANTIGENS AND
;    TITLE OF INVENTION:  THEIR USE IN DIAGNOSTIC AND THERAPEUTIC
;    TITLE OF INVENTION:  METHODS
```

```
;    NUMBER OF SEQUENCES:  43
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MORGAN & FINNEGAN
;      STREET:  345 PARK AVENUE
;      CITY:  NEW YORK
;      STATE:  NEW YORK
;      COUNTRY:  USA
;      ZIP:  10154
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  FLOPPY DISK
;      COMPUTER:  IBM PC COMPATIBLE
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  ASCII
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/231,565A
;      FILING DATE:  22-APR-1994
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  CAROL M. GRUPPI
;      REGISTRATION NUMBER:  37,341
;      REFERENCE/DOCKET NUMBER:  2026-4124
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 758-4800
;      TELEFAX:  (212) 751-6849
;      TELEX:  421792
;  INFORMATION FOR SEQ ID NO:  13:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9
;      TYPE:  amino acid
;      STRANDEDNESS:  Unknown
;      TOPOLOGY:  Unknown
;    MOLECULE TYPE:  Peptide
US-08-231-565A-13

  Query Match             47.1%;  Score 24;  DB 2;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 1.5e+05;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       4 PNAP 7
           ||||
Db       4 PNAP 7


RESULT  13
US-09-007-961-13
; Sequence 13, Application US/09007961
; Patent No. 5994523
;  GENERAL INFORMATION:
;    APPLICANT:  KAWAKAMI, YUTAKA; ROSENBERG,
;    APPLICANT:  STEVEN A.
;    TITLE OF INVENTION:  MELANOMA ANTIGENS AND
;    TITLE OF INVENTION:  THEIR USE IN DIAGNOSTIC AND THERAPEUTIC
;    TITLE OF INVENTION:  METHODS
;    NUMBER OF SEQUENCES:  43
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MORGAN & FINNEGAN
;      STREET:  345 PARK AVENUE
;      CITY:  NEW YORK
;      STATE:  NEW YORK
;      COUNTRY:  USA
;      ZIP:  10154
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  FLOPPY DISK
;      COMPUTER:  IBM PC COMPATIBLE
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  ASCII
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/007,961
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/231,565
```

```
;       FILING DATE:  22-APR-1994
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  CAROL M. GRUPPI
;       REGISTRATION NUMBER:  37,341
;       REFERENCE/DOCKET NUMBER:  2026-4124
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (212) 758-4800
;       TELEFAX:  (212) 751-6849
;       TELEX:  421792
;   INFORMATION FOR SEQ ID NO:  13:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  9
;       TYPE:  amino acid
;       STRANDEDNESS:  Unknown
;       TOPOLOGY:  Unknown
;     MOLECULE TYPE:  Peptide
US-09-007-961-13


  Query Match             47.1%;  Score 24;  DB 2;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 1.5e+05;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      4 PNAP 7
          ||||
Db      4 PNAP 7


RESULT  14
US-08-159-339A-393
; Sequence 393, Application US/08159339A
; Patent No. 6037135
;  GENERAL INFORMATION:
;     APPLICANT:  Kubo, Ralph T.
;     APPLICANT:  Grey, Howard M.
;     APPLICANT:  Sette, Alessandro
;     APPLICANT:  Celis, Esteban
;     TITLE OF INVENTION:  HLA Binding peptides and Their
;     TITLE OF INVENTION:  Uses
;     NUMBER OF SEQUENCES:  1254
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Townsend and Townsend and Crew LLP
;       STREET:  Two Embarcadero Center, Eighth Floor
;       CITY:  San Francisco
;       STATE:  CA
;       COUNTRY:  USA
;       ZIP:  94111-3834
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Diskette
;       COMPUTER:  IBM Compatible
;       OPERATING SYSTEM:  DOS
;       SOFTWARE:  FastSEQ for Windows Version 2.0
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/159,339A
;       FILING DATE:  29-NOV-1993
;       CLASSIFICATION:  424
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 07/926,666
;       FILING DATE:  07-AUG-1992
;       APPLICATION NUMBER:  US 08/027,746
;       FILING DATE:  05-MAR-1993
;       APPLICATION NUMBER:  US 08/103,396
;       FILING DATE:  06-AUG-1993
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Weber, Ellen Lauver
;       REGISTRATION NUMBER:  32,762
;       REFERENCE/DOCKET NUMBER:  018623-005030US
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (415) 576-0200
;       TELEFAX:  (415) 576-0300
;       TELEX:
;   INFORMATION FOR SEQ ID NO:  393:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  9 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
US-08-159-339A-393

  Query Match             47.1%;  Score 24;  DB 3;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 1.5e+05;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      4 PNAP 7
          ||||
Db      5 PNAP 8


RESULT  15
US-09-162-368B-24
; Sequence 24, Application US/09162368B
; Patent No. 6083703
;  GENERAL INFORMATION:
;     APPLICANT:  WANG, R.F.; ROSENBERG, S. A.
;     TITLE OF INVENTION:  IDENTIFICATION OF TRP-2 AS
;     TITLE OF INVENTION:  A HUMAN TUMOR ANTIGEN RECOGNIZED BY CYTOTOXIC T
;     TITLE OF INVENTION:  LYMPHOCYTES
;     NUMBER OF SEQUENCES:  31
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  MORGAN & FINNEGAN, L.L.P.
;       STREET:  345 PARK AVENUE
;       CITY:  NEW YORK
;       STATE:  NEW YORK
;       COUNTRY:  USA
;       ZIP:  10154
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  FLOPPY DISK
;       COMPUTER:  IBM PC COMPATIBLE
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  MICROSOFT WORD 97
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/09/162,368B
;       FILING DATE:  28-SEPT-1998
;       CLASSIFICATION:  530
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  08/725,736
;       FILING DATE:  04-OCT-1996
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  08/599,602
;       FILING DATE:  09-FEB-1996
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  KATHRYN M. BROWN
;       REGISTRATION NUMBER:  34,556
;       REFERENCE/DOCKET NUMBER:  2026-4243US1
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (212) 758-4800
;       TELEFAX:  (212) 751-6849
;       TELEX:  421792
;   INFORMATION FOR SEQ ID NO:  24:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  9
;       TYPE:  AMINO ACID
;       STRANDEDNESS:  UNKNOWN
;       TOPOLOGY:  UNKNOWN
;     MOLECULE TYPE:
;       DESCRIPTION:  PEPTIDE
;     FEATURE:
;       NAME/KEY:
;       LOCATION:
;       IDENTIFICATION METHOD:
;       OTHER INFORMATION:
US-09-162-368B-24
```

```
  Query Match             47.1%;  Score 24;  DB 3;  Length 9;
  Best Local Similarity   50.0%;  Pred. No. 1.5e+05;
  Matches    4;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 RMFPNAPY 8
           |: |  ||
Db       1 RLGPGRPY 8


Search completed: October 23, 2001, 13:29:21
Job time: 389 sec
```

```
                    GenCore version 4.5
             Copyright (c) 1993 - 2000  Compugen Ltd.

OM protein - protein search, using sw model

Run on:         October 23, 2001, 13:27:58 ; Search time 26.91 Seconds
                                             (without alignments)
                                             17.233 Million cell updates/sec

Title:          US-09-625-963-1
Perfect score:  51
Sequence:       1 RMFPNAPYL 9

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched:       339663 seqs, 51527020 residues

Total number of hits satisfying chosen parameters:      53587

Minimum DB seq length: 0
Maximum DB seq length: 9

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       Pending_Patents_AA_New:*
                1:  /cgn2_6/ptodata/2/paa/PCT_NEW_COMB.pep:*
                2:  /cgn2_6/ptodata/2/paa/US06_NEW_COMB.pep:*
                3:  /cgn2_6/ptodata/2/paa/US07_NEW_COMB.pep:*
                4:  /cgn2_6/ptodata/2/paa/US08_NEW_COMB.pep:*
                5:  /cgn2_6/ptodata/2/paa/US09_NEW_COMB.pep:*
                6:  /cgn2_6/ptodata/2/paa/US60_NEW_COMB.pep:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                              SUMMARIES
                %
Result          Query
  No.   Score   Match Length DB  ID                         Description
-------------------------------------------------------------------------------
     1     51   100.0      9  5  US-09-938-864-185           Sequence 185, App
     2     51   100.0      9  5  US-09-938-864-293           Sequence 293, App
     3     47    92.2      9  5  US-09-938-864-38            Sequence 38, Appl
     4     46    90.2      9  5  US-09-938-864-324           Sequence 324, App
     5     41    80.4      9  5  US-09-938-864-79            Sequence 79, Appl
     6     41    80.4      9  5  US-09-938-864-267           Sequence 267, App
     7     40    78.4      9  5  US-09-938-864-323           Sequence 323, App
     8     35    68.6      9  5  US-09-938-864-325           Sequence 325, App
     9     33    64.7      9  5  US-09-938-864-98            Sequence 98, Appl
    10     33    64.7      9  5  US-09-938-864-269           Sequence 269, App
    11     29    56.9      9  5  US-09-938-864-212           Sequence 212, App
    12     28    54.9      9  5  US-09-938-864-144           Sequence 144, App
    13     28    54.9      9  5  US-09-938-864-282           Sequence 282, App
    14     25    49.0      9  5  US-09-884-441-436           Sequence 436, App
    15     25    49.0      9  5  US-09-884-441-447           Sequence 447, App
    16     25    49.0      9  5  US-09-667-857-436           Sequence 436, App
    17     25    49.0      9  5  US-09-667-857-447           Sequence 447, App
    18     25    49.0      9  5  US-09-907-969-436           Sequence 436, App
    19     25    49.0      9  5  US-09-907-969-447           Sequence 447, App
    20     24    47.1      8  5  US-09-458-298-792           Sequence 792, App
    21     23    45.1      9  5  US-09-938-864-322           Sequence 322, App
    22     22    43.1      7  5  US-09-571-407-232           Sequence 232, App
    23     22    43.1      7  5  US-09-571-407-234           Sequence 234, App
    24     22    43.1      8  5  US-09-571-407-225           Sequence 225, App
    25     22    43.1      8  5  US-09-571-407-227           Sequence 227, App
    26     22    43.1      8  5  US-09-166-845B-21           Sequence 21, Appl
    27     22    43.1      8  5  US-09-757-933B-8            Sequence 8, Appli
    28     22    43.1      9  5  US-09-571-407-224           Sequence 224, App
    29     22    43.1      9  5  US-09-571-407-226           Sequence 226, App
    30     22    43.1      9  5  US-09-938-864-326           Sequence 326, App
    31     21    41.2      8  5  US-09-399-220B-36           Sequence 36, Appl
    32     21    41.2      8  5  US-09-399-220B-55           Sequence 55, Appl
    33     21    41.2      8  5  US-09-399-220B-58           Sequence 58, Appl
    34     20    39.2      7  5  US-09-687-156A-11           Sequence 11, Appl
    35     20    39.2      9  5  US-09-832-723-99            Sequence 99, Appl
    36     19    37.3      4  5  US-09-786-130-19            Sequence 19, Appl
    37     19    37.3      4  5  US-09-744-931-35            Sequence 35, Appl
    38     19    37.3      5  5  US-09-595-682B-6            Sequence 6, Appli
    39     19    37.3      6  5  US-09-801-852A-9            Sequence 9, Appli
    40     19    37.3      6  5  US-09-645-793A-32           Sequence 32, Appl
    41     19    37.3      7  5  US-09-812-283-35            Sequence 35, Appl
    42     19    37.3      7  5  US-09-832-723-7             Sequence 7, Appli
    43     19    37.3      8  5  US-09-786-569-4             Sequence 4, Appli
    44     19    37.3      9  4  US-08-821-828-5             Sequence 5, Appli
    45     19    37.3      9  5  US-09-519-658A-56           Sequence 56, Appl

                              ALIGNMENTS

RESULT   1
US-09-938-864-185
; Sequence 185, Application US/09938864
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Smithgall, Molly
;  APPLICANT:  Moulton, Gus
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Sleath, Paul R.
;  APPLICANT:  Mossman, Sally
;  APPLICANT:  Evans, Lawrence
;  APPLICANT:  Spies, A. Gregory
;  APPLICANT:  Boydston, Jeremy
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1
;  TITLE OF INVENTION:  SPECIFIC IMMUNOTHERAPY
;  FILE REFERENCE: 210121.465C5
;  CURRENT APPLICATION NUMBER: US/09/938,864
;  CURRENT FILING DATE:  2001-08-24
;  NUMBER OF SEQ ID NOS: 413
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 185
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Homo sapien
US-09-938-864-185


  Query Match             100.0%;  Score 51;  DB 5;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.9e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 RMFPNAPYL 9
          |||||||||
Db      1 RMFPNAPYL 9


RESULT   2
US-09-938-864-293
; Sequence 293, Application US/09938864
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Smithgall, Molly
;  APPLICANT:  Moulton, Gus
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Sleath, Paul R.
;  APPLICANT:  Mossman, Sally
;  APPLICANT:  Evans, Lawrence
;  APPLICANT:  Spies, A. Gregory
```

```
;  APPLICANT:  Boydston, Jeremy
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1
;  TITLE OF INVENTION:  SPECIFIC IMMUNOTHERAPY
;  FILE REFERENCE: 210121.465C5
;  CURRENT APPLICATION NUMBER: US/09/938,864
;  CURRENT FILING DATE:  2001-08-24
;  NUMBER OF SEQ ID NOS: 413
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 293
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-938-864-293


  Query Match             100.0%;  Score 51;  DB 5;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.9e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RMFPNAPYL 9
           |||||||||
Db       1 RMFPNAPYL 9


RESULT   3
US-09-938-864-38
; Sequence 38, Application US/09938864
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Smithgall, Molly
;  APPLICANT:  Moulton, Gus
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Sleath, Paul R.
;  APPLICANT:  Mossman, Sally
;  APPLICANT:  Evans, Lawrence
;  APPLICANT:  Spies, A. Gregory
;  APPLICANT:  Boydston, Jeremy
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1
;  TITLE OF INVENTION:  SPECIFIC IMMUNOTHERAPY
;  FILE REFERENCE: 210121.465C5
;  CURRENT APPLICATION NUMBER: US/09/938,864
;  CURRENT FILING DATE:  2001-08-24
;  NUMBER OF SEQ ID NOS: 413
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 38
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Homo sapien
US-09-938-864-38


  Query Match             92.2%;  Score 47;  DB 5;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.9e+05;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RMFPNAPY 8
           ||||||||
Db       2 RMFPNAPY 9


RESULT   4
US-09-938-864-324
; Sequence 324, Application US/09938864
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Smithgall, Molly
;  APPLICANT:  Moulton, Gus
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Sleath, Paul R.
;  APPLICANT:  Mossman, Sally
;  APPLICANT:  Evans, Lawrence
;  APPLICANT:  Spies, A. Gregory
;  APPLICANT:  Boydston, Jeremy
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1
;  TITLE OF INVENTION:  SPECIFIC IMMUNOTHERAPY
;  FILE REFERENCE: 210121.465C5
;  CURRENT APPLICATION NUMBER: US/09/938,864
;  CURRENT FILING DATE:  2001-08-24
;  NUMBER OF SEQ ID NOS: 413
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 324
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Homo sapien and Mus musculus
US-09-938-864-324


  Query Match             90.2%;  Score 46;  DB 5;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.9e+05;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2 MFPNAPYL 9
           ||||||||
Db       1 MFPNAPYL 8


RESULT   5
US-09-938-864-79
; Sequence 79, Application US/09938864
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Smithgall, Molly
;  APPLICANT:  Moulton, Gus
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Sleath, Paul R.
;  APPLICANT:  Mossman, Sally
;  APPLICANT:  Evans, Lawrence
;  APPLICANT:  Spies, A. Gregory
;  APPLICANT:  Boydston, Jeremy
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1
;  TITLE OF INVENTION:  SPECIFIC IMMUNOTHERAPY
;  FILE REFERENCE: 210121.465C5
;  CURRENT APPLICATION NUMBER: US/09/938,864
;  CURRENT FILING DATE:  2001-08-24
;  NUMBER OF SEQ ID NOS: 413
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 79
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Homo sapien
US-09-938-864-79


  Query Match             80.4%;  Score 41;  DB 5;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.9e+05;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3 FPNAPYL 9
           |||||||
Db       1 FPNAPYL 7


RESULT   6
US-09-938-864-267
; Sequence 267, Application US/09938864
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Smithgall, Molly
;  APPLICANT:  Moulton, Gus
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Sleath, Paul R.
;  APPLICANT:  Mossman, Sally
```

```
;  APPLICANT:  Evans, Lawrence
;  APPLICANT:  Spies, A. Gregory
;  APPLICANT:  Boydston, Jeremy
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1
;  TITLE OF INVENTION:  SPECIFIC IMMUNOTHERAPY
;  FILE REFERENCE: 210121.465C5
;  CURRENT APPLICATION NUMBER: US/09/938,864
;  CURRENT FILING DATE:  2001-08-24
;  NUMBER OF SEQ ID NOS: 413
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 267
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-938-864-267

  Query Match            80.4%;  Score 41;  DB 5;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.9e+05;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3 FPNAPYL 9
           |||||||
Db       1 FPNAPYL 7

RESULT   7
US-09-938-864-323
; Sequence 323, Application US/09938864
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Smithgall, Molly
;  APPLICANT:  Moulton, Gus
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Sleath, Paul R.
;  APPLICANT:  Mossman, Sally
;  APPLICANT:  Evans, Lawrence
;  APPLICANT:  Spies, A. Gregory
;  APPLICANT:  Boydston, Jeremy
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1
;  TITLE OF INVENTION:  SPECIFIC IMMUNOTHERAPY
;  FILE REFERENCE: 210121.465C5
;  CURRENT APPLICATION NUMBER: US/09/938,864
;  CURRENT FILING DATE:  2001-08-24
;  NUMBER OF SEQ ID NOS: 413
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 323
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Homo sapien and Mus musculus
US-09-938-864-323

  Query Match            78.4%;  Score 40;  DB 5;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.9e+05;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RMFPNAP 7
           |||||||
Db       3 RMFPNAP 9

RESULT   8
US-09-938-864-325
; Sequence 325, Application US/09938864
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Smithgall, Molly
;  APPLICANT:  Moulton, Gus
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Sleath, Paul R.
```

```
;  APPLICANT:  Mossman, Sally
;  APPLICANT:  Evans, Lawrence
;  APPLICANT:  Spies, A. Gregory
;  APPLICANT:  Boydston, Jeremy
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1
;  TITLE OF INVENTION:  SPECIFIC IMMUNOTHERAPY
;  FILE REFERENCE: 210121.465C5
;  CURRENT APPLICATION NUMBER: US/09/938,864
;  CURRENT FILING DATE:  2001-08-24
;  NUMBER OF SEQ ID NOS: 413
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 325
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Homo sapien and Mus musculus
US-09-938-864-325

  Query Match            68.6%;  Score 35;  DB 5;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.9e+05;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       4 PNAPYL 9
           ||||||
Db       1 PNAPYL 6

RESULT   9
US-09-938-864-98
; Sequence 98, Application US/09938864
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Smithgall, Molly
;  APPLICANT:  Moulton, Gus
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Sleath, Paul R.
;  APPLICANT:  Mossman, Sally
;  APPLICANT:  Evans, Lawrence
;  APPLICANT:  Spies, A. Gregory
;  APPLICANT:  Boydston, Jeremy
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1
;  TITLE OF INVENTION:  SPECIFIC IMMUNOTHERAPY
;  FILE REFERENCE: 210121.465C5
;  CURRENT APPLICATION NUMBER: US/09/938,864
;  CURRENT FILING DATE:  2001-08-24
;  NUMBER OF SEQ ID NOS: 413
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 98
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Homo sapien
US-09-938-864-98

  Query Match            64.7%;  Score 33;  DB 5;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.9e+05;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RMFPNA 6
           ||||||
Db       4 RMFPNA 9

RESULT  10
US-09-938-864-269
; Sequence 269, Application US/09938864
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Smithgall, Molly
;  APPLICANT:  Moulton, Gus
;  APPLICANT:  Vedvick, Thomas S.
```

```
;  APPLICANT:  Sleath, Paul R.
;  APPLICANT:  Mossman, Sally
;  APPLICANT:  Evans, Lawrence
;  APPLICANT:  Spies, A. Gregory
;  APPLICANT:  Boydston, Jeremy
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1
;  TITLE OF INVENTION:  SPECIFIC IMMUNOTHERAPY
;  FILE REFERENCE: 210121.465C5
;  CURRENT APPLICATION NUMBER: US/09/938,864
;  CURRENT FILING DATE:  2001-08-24
;  NUMBER OF SEQ ID NOS: 413
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 269
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-938-864-269

  Query Match             64.7%;  Score 33;  DB 5;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.9e+05;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RMFPNA 6
              ||||||
Db          4 RMFPNA 9

RESULT  11
US-09-938-864-212
; Sequence 212, Application US/09938864
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Smithgall, Molly
;  APPLICANT:  Moulton, Gus
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Sleath, Paul R.
;  APPLICANT:  Mossman, Sally
;  APPLICANT:  Evans, Lawrence
;  APPLICANT:  Spies, A. Gregory
;  APPLICANT:  Boydston, Jeremy
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1
;  TITLE OF INVENTION:  SPECIFIC IMMUNOTHERAPY
;  FILE REFERENCE: 210121.465C5
;  CURRENT APPLICATION NUMBER: US/09/938,864
;  CURRENT FILING DATE:  2001-08-24
;  NUMBER OF SEQ ID NOS: 413
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 212
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Homo sapien
US-09-938-864-212

  Query Match             56.9%;  Score 29;  DB 5;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.9e+05;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RMFPN 5
              |||||
Db          5 RMFPN 9

RESULT  12
US-09-938-864-144
; Sequence 144, Application US/09938864
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Smithgall, Molly
;  APPLICANT:  Moulton, Gus
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Sleath, Paul R.
;  APPLICANT:  Mossman, Sally
;  APPLICANT:  Evans, Lawrence
;  APPLICANT:  Spies, A. Gregory
;  APPLICANT:  Boydston, Jeremy
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1
;  TITLE OF INVENTION:  SPECIFIC IMMUNOTHERAPY
;  FILE REFERENCE: 210121.465C5
;  CURRENT APPLICATION NUMBER: US/09/938,864
;  CURRENT FILING DATE:  2001-08-24
;  NUMBER OF SEQ ID NOS: 413
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 144
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Homo sapien
US-09-938-864-144

  Query Match             54.9%;  Score 28;  DB 5;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.9e+05;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          5 NAPYL 9
              |||||
Db          1 NAPYL 5

RESULT  13
US-09-938-864-282
; Sequence 282, Application US/09938864
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Smithgall, Molly
;  APPLICANT:  Moulton, Gus
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Sleath, Paul R.
;  APPLICANT:  Mossman, Sally
;  APPLICANT:  Evans, Lawrence
;  APPLICANT:  Spies, A. Gregory
;  APPLICANT:  Boydston, Jeremy
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1
;  TITLE OF INVENTION:  SPECIFIC IMMUNOTHERAPY
;  FILE REFERENCE: 210121.465C5
;  CURRENT APPLICATION NUMBER: US/09/938,864
;  CURRENT FILING DATE:  2001-08-24
;  NUMBER OF SEQ ID NOS: 413
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 282
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-938-864-282

  Query Match             54.9%;  Score 28;  DB 5;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.9e+05;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          5 NAPYL 9
              |||||
Db          1 NAPYL 5

RESULT  14
US-09-884-441-436
; Sequence 436, Application US/09884441
; GENERAL INFORMATION:
;  APPLICANT: Algate, Paul A.
;  APPLICANT:  Carter, Darrick
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
```

```
;  TITLE OF INVENTION:  DIAGNOSIS OF OVARIAN CANCER
;  FILE REFERENCE: 210121.462C7
;  CURRENT APPLICATION NUMBER: US/09/884,441
;  CURRENT FILING DATE:  2001-06-18
;  NUMBER OF SEQ ID NOS: 489
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 436
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-884-441-436


  Query Match             49.0%;  Score 25;  DB 5;  Length 9;
  Best Local Similarity   50.0%;  Pred. No. 2.9e+05;
  Matches    4;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 MFPNAPYL 9
              : | :|||
Db          2 LLPLSPYL 9


RESULT  15
US-09-884-441-447
; Sequence 447, Application US/09884441
; GENERAL INFORMATION:
;  APPLICANT: Algate, Paul A.
;  APPLICANT:  Carter, Darrick
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF OVARIAN CANCER
;  FILE REFERENCE: 210121.462C7
;  CURRENT APPLICATION NUMBER: US/09/884,441
;  CURRENT FILING DATE:  2001-06-18
;  NUMBER OF SEQ ID NOS: 489
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 447
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-884-441-447


  Query Match             49.0%;  Score 25;  DB 5;  Length 9;
  Best Local Similarity   50.0%;  Pred. No. 2.9e+05;
  Matches    4;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 MFPNAPYL 9
              : | :|||
Db          1 LLPLSPYL 8


Search completed: October 23, 2001, 13:30:23
Job time: 145 sec
```

```
                        GenCore version 4.5
                Copyright (c) 1993 - 2000  Compugen Ltd.


OM protein - protein search, using sw model

Run on:         October 23, 2001, 13:28:23 ; Search time 155.83 Seconds
                                             (without alignments)
                                             12.813 Million cell updates/sec

Title:          US-09-625-963-1
Perfect score:  51
Sequence:       1 RMFPNAPYL 9

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched:       2803329 seqs, 221847457 residues

Total number of hits satisfying chosen parameters:      162356

Minimum DB seq length: 0
Maximum DB seq length: 9

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :      Pending_Patents_AA_Main:*
              1:  /cgn2_6/ptodata/2/paa/PCTUS_COMB.pep:*
              2:  /cgn2_6/ptodata/2/paa/US06_COMB.pep:*
              3:  /cgn2_6/ptodata/2/paa/US07_COMB.pep:*
              4:  /cgn2_6/ptodata/2/paa/US080_COMB.pep:*
              5:  /cgn2_6/ptodata/2/paa/US081_COMB.pep:*
              6:  /cgn2_6/ptodata/2/paa/US082_COMB.pep:*
              7:  /cgn2_6/ptodata/2/paa/US083_COMB.pep:*
              8:  /cgn2_6/ptodata/2/paa/US084_COMB.pep:*
              9:  /cgn2_6/ptodata/2/paa/US085_COMB.pep:*
             10:  /cgn2_6/ptodata/2/paa/US086_COMB.pep:*
             11:  /cgn2_6/ptodata/2/paa/US087_COMB.pep:*
             12:  /cgn2_6/ptodata/2/paa/US088_COMB.pep:*
             13:  /cgn2_6/ptodata/2/paa/US089_COMB.pep:*
             14:  /cgn2_6/ptodata/2/paa/US090_COMB.pep:*
             15:  /cgn2_6/ptodata/2/paa/US091_COMB.pep:*
             16:  /cgn2_6/ptodata/2/paa/US092_COMB.pep:*
             17:  /cgn2_6/ptodata/2/paa/US093_COMB.pep:*
             18:  /cgn2_6/ptodata/2/paa/US094_COMB.pep:*
             19:  /cgn2_6/ptodata/2/paa/US095_COMB.pep:*
             20:  /cgn2_6/ptodata/2/paa/US096_COMB.pep:*
             21:  /cgn2_6/ptodata/2/paa/US097_COMB.pep:*
             22:  /cgn2_6/ptodata/2/paa/US098_COMB.pep:*
             23:  /cgn2_6/ptodata/2/paa/US60_COMB.pep:*

     Pred. No. is the number of results predicted by chance to have a
     score greater than or equal to the score of the result being printed,
     and is derived by analysis of the total score distribution.

                                SUMMARIES
                  %
Result           Query
  No.   Score   Match Length DB  ID                    Description
--------------------------------------------------------------------------
    1     51   100.0      9  15  US-09-164-223-185       Sequence 185, App
    2     51   100.0      9  15  US-09-164-223-293       Sequence 293, App
    3     51   100.0      9  15  US-09-164-223-185       Sequence 185, App
    4     51   100.0      9  15  US-09-164-223-293       Sequence 293, App
    5     51   100.0      9  15  US-09-164-223-185       Sequence 185, App
    6     51   100.0      9  15  US-09-164-223-293       Sequence 293, App
    7     51   100.0      9  16  US-09-276-484-185       Sequence 185, App
    8     51   100.0      9  16  US-09-276-484-293       Sequence 293, App
    9     51   100.0      9  16  US-09-276-484A-185      Sequence 185, App
   10     51   100.0      9  16  US-09-276-484A-293      Sequence 293, App
   11     51   100.0      9  19  US-09-523-419-185       Sequence 185, App
   12     51   100.0      9  19  US-09-523-419-293       Sequence 293, App
   13     51   100.0      9  20  US-09-625-963-1         Sequence 1, Appli
   14     51   100.0      9  20  US-09-679-339-185       Sequence 185, App
   15     51   100.0      9  20  US-09-679-339-293       Sequence 293, App
   16     51   100.0      9  20  US-09-684-361-185       Sequence 185, App
   17     51   100.0      9  20  US-09-684-361-293       Sequence 293, App
   18     51   100.0      9  20  US-09-685-830-185       Sequence 185, App
   19     51   100.0      9  20  US-09-685-830-293       Sequence 293, App
   20     51   100.0      9  21  US-09-785-019-185       Sequence 185, App
   21     51   100.0      9  21  US-09-785-019-293       Sequence 293, App
   22     51   100.0      9  21  US-09-791-477-185       Sequence 185, App
   23     51   100.0      9  21  US-09-791-477-293       Sequence 293, App
   24     51   100.0      9  22  US-09-872-832-46        Sequence 46, Appl
   25     47    92.2      9  15  US-09-164-223-38        Sequence 38, Appl
   26     47    92.2      9  15  US-09-164-223-38        Sequence 38, Appl
   27     47    92.2      9  15  US-09-164-223-38        Sequence 38, Appl
   28     47    92.2      9  16  US-09-276-484-38        Sequence 38, Appl
   29     47    92.2      9  16  US-09-276-484A-38       Sequence 38, Appl
   30     47    92.2      9  19  US-09-523-419-38        Sequence 38, Appl
   31     47    92.2      9  20  US-09-679-339-38        Sequence 38, Appl
   32     47    92.2      9  20  US-09-684-361-38        Sequence 38, Appl
   33     47    92.2      9  20  US-09-685-830-38        Sequence 38, Appl
   34     47    92.2      9  21  US-09-785-019-38        Sequence 38, Appl
   35     47    92.2      9  21  US-09-791-477-38        Sequence 38, Appl
   36     46    90.2      9  16  US-09-276-484-324       Sequence 324, App
   37     46    90.2      9  16  US-09-276-484A-324      Sequence 324, App
   38     46    90.2      9  19  US-09-523-419-324       Sequence 324, App
   39     46    90.2      9  20  US-09-679-339-324       Sequence 324, App
   40     46    90.2      9  20  US-09-684-361-324       Sequence 324, App
   41     46    90.2      9  20  US-09-685-830-324       Sequence 324, App
   42     46    90.2      9  21  US-09-785-019-324       Sequence 324, App
   43     46    90.2      9  21  US-09-791-477-324       Sequence 324, App
   44     41    80.4      9  15  US-09-164-223-79        Sequence 79, Appl
   45     41    80.4      9  15  US-09-164-223-267       Sequence 267, App


                                ALIGNMENTS


RESULT   1
US-09-164-223-185
; Sequence 185, Application US/09164223
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
;  APPLICANT: Cheever, Martin A.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1 SPECIFIC IMMUNOTHERAPY
;  FILE REFERENCE: 210121.465
;  CURRENT APPLICATION NUMBER: US/09/164,223
;  CURRENT FILING DATE: 1998-09-30
;  NUMBER OF SEQ ID NOS: 320
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 185
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-164-223-185

  Query Match             100.0%;  Score 51;  DB 15;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.6e+06;
  Matches    9; Conservative    0; Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RMFPNAPYL 9
           |||||||||
Db       1 RMFPNAPYL 9


RESULT   2
US-09-164-223-293
; Sequence 293, Application US/09164223
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
```

```
;  APPLICANT: Cheever, Martin A.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1 SPECIFIC IMMUNOTHERAPY
;  FILE REFERENCE: 210121.465
;  CURRENT APPLICATION NUMBER: US/09/164,223
;  CURRENT FILING DATE: 1998-09-30
;  NUMBER OF SEQ ID NOS: 320
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 293
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-164-223-293


  Query Match             100.0%;  Score 51;  DB 15;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.6e+06;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 RMFPNAPYL 9
            |||||||||
Db        1 RMFPNAPYL 9

RESULT   3
US-09-164-223-185
; Sequence 185, Application US/09164223A
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
;  APPLICANT: Cheever, Martin A.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1 SPECIFIC IMMUNOTHERAPY
;  FILE REFERENCE: 210121.465
;  CURRENT APPLICATION NUMBER: US/09/164,223A
;  CURRENT FILING DATE: 1998-09-30
;  NUMBER OF SEQ ID NOS: 320
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 185
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-164-223-185


  Query Match             100.0%;  Score 51;  DB 15;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.6e+06;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 RMFPNAPYL 9
            |||||||||
Db        1 RMFPNAPYL 9

RESULT   4
US-09-164-223-293
; Sequence 293, Application US/09164223A
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
;  APPLICANT: Cheever, Martin A.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1 SPECIFIC IMMUNOTHERAPY
;  FILE REFERENCE: 210121.465
;  CURRENT APPLICATION NUMBER: US/09/164,223A
;  CURRENT FILING DATE: 1998-09-30
;  NUMBER OF SEQ ID NOS: 320
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 293
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-164-223-293


  Query Match             100.0%;  Score 51;  DB 15;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.6e+06;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy        1 RMFPNAPYL 9
            |||||||||
Db        1 RMFPNAPYL 9

RESULT   5
US-09-164-223-185
; Sequence 185, Application US/09164223B
; GENERAL INFORMATION:
;  APPLICANT:  Gaiger, Alexander
;  APPLICANT:  Cheever, Martin A.
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR WT1 SPECIFIC IMMUNOTHERAPY
;  FILE REFERENCE:  210121.465
;  CURRENT APPLICATION NUMBER: US/09/164,223B
;  CURRENT FILING DATE:  1998-09-30
;  NUMBER OF SEQ ID NOS:  320
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 185
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-164-223-185


  Query Match             100.0%;  Score 51;  DB 15;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.6e+06;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 RMFPNAPYL 9
            |||||||||
Db        1 RMFPNAPYL 9


RESULT   6
US-09-164-223-293
; Sequence 293, Application US/09164223B
; GENERAL INFORMATION:
;  APPLICANT:  Gaiger, Alexander
;  APPLICANT:  Cheever, Martin A.
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR WT1 SPECIFIC IMMUNOTHERAPY
;  FILE REFERENCE:  210121.465
;  CURRENT APPLICATION NUMBER: US/09/164,223B
;  CURRENT FILING DATE:  1998-09-30
;  NUMBER OF SEQ ID NOS:  320
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 293
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-164-223-293


  Query Match             100.0%;  Score 51;  DB 15;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.6e+06;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 RMFPNAPYL 9
            |||||||||
Db        1 RMFPNAPYL 9


RESULT   7
US-09-276-484-185
; Sequence 185, Application US/09276484
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
;  APPLICANT: Cheever, Martin A.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1
;  TITLE OF INVENTION: SPECIFIC IMMUNOTHERAPY
;  FILE REFERENCE: 210121.465C1
;  CURRENT APPLICATION NUMBER: US/09/276,484
;  CURRENT FILING DATE: 1999-03-25
```

```
;  NUMBER OF SEQ ID NOS: 326
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 185
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Homo sapien
US-09-276-484-185


  Query Match             100.0%;  Score 51;  DB 16;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.6e+06;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 RMFPNAPYL 9
            |||||||||
Db        1 RMFPNAPYL 9


RESULT   8
US-09-276-484-293
; Sequence 293, Application US/09276484
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
;  APPLICANT: Cheever, Martin A.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1
;  TITLE OF INVENTION: SPECIFIC IMMUNOTHERAPY
;  FILE REFERENCE: 210121.465C1
;  CURRENT APPLICATION NUMBER: US/09/276,484
;  CURRENT FILING DATE: 1999-03-25
;  NUMBER OF SEQ ID NOS: 326
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 293
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-276-484-293


  Query Match             100.0%;  Score 51;  DB 16;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.6e+06;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 RMFPNAPYL 9
            |||||||||
Db        1 RMFPNAPYL 9


RESULT   9
US-09-276-484A-185
; Sequence 185, Application US/09276484A
; GENERAL INFORMATION:
;  APPLICANT:  Gaiger, Alexander
;  APPLICANT:  Cheever, Martin A.
;  TITLE OF INVENTION:  COMPOSITONS AND METHODS FOR WT1 SPECIFIC
;  TITLE OF INVENTION:  IMMUNOTHERAPY
;  FILE REFERENCE:  210121.465C1
;  CURRENT APPLICATION NUMBER: US/09/276,484A
;  CURRENT FILING DATE:  1999-03-25
;  PRIOR APPLICATION NUMBER:  US 09/164,223
;  PRIOR FILING DATE:  1999-09-30
;  NUMBER OF SEQ ID NOS:  326
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 185
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Homo sapien
US-09-276-484A-185


  Query Match             100.0%;  Score 51;  DB 16;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.6e+06;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 RMFPNAPYL 9
            |||||||||
Db        1 RMFPNAPYL 9


RESULT  10
US-09-276-484A-293
; Sequence 293, Application US/09276484A
; GENERAL INFORMATION:
;  APPLICANT:  Gaiger, Alexander
;  APPLICANT:  Cheever, Martin A.
;  TITLE OF INVENTION:  COMPOSITONS AND METHODS FOR WT1 SPECIFIC
;  TITLE OF INVENTION:  IMMUNOTHERAPY
;  FILE REFERENCE:  210121.465C1
;  CURRENT APPLICATION NUMBER: US/09/276,484A
;  CURRENT FILING DATE:  1999-03-25
;  PRIOR APPLICATION NUMBER:  US 09/164,223
;  PRIOR FILING DATE:  1999-09-30
;  NUMBER OF SEQ ID NOS:  326
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 293
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-276-484A-293


  Query Match             100.0%;  Score 51;  DB 16;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.6e+06;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 RMFPNAPYL 9
            |||||||||
Db        1 RMFPNAPYL 9


RESULT  11
US-09-523-419-185
; Sequence 185, Application US/09523419
; GENERAL INFORMATION:
;  APPLICANT: Hendrickson, Ronald C.
;  APPLICANT: Probst, Peter
;  APPLICANT: Reed, Steven G.
;  TITLE OF INVENTION: METHODS FOR THE DIAGNOSIS AND THERAPY
;  TITLE OF INVENTION: OF CHLAMYDIA INFECTION
;  FILE REFERENCE: 210121.492
;  CURRENT APPLICATION NUMBER: US/09/523,419
;  CURRENT FILING DATE: 2000-03-10
;  NUMBER OF SEQ ID NOS: 326
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 185
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Homo sapien
US-09-523-419-185


  Query Match             100.0%;  Score 51;  DB 19;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.6e+06;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 RMFPNAPYL 9
            |||||||||
Db        1 RMFPNAPYL 9


RESULT  12
US-09-523-419-293
; Sequence 293, Application US/09523419
; GENERAL INFORMATION:
;  APPLICANT: Hendrickson, Ronald C.
;  APPLICANT: Probst, Peter
;  APPLICANT: Reed, Steven G.
```

```
;  TITLE OF INVENTION: METHODS FOR THE DIAGNOSIS AND THERAPY
;  TITLE OF INVENTION: OF CHLAMYDIA INFECTION
;  FILE REFERENCE: 210121.492
;  CURRENT APPLICATION NUMBER: US/09/523,419
;  CURRENT FILING DATE: 2000-03-10
;  NUMBER OF SEQ ID NOS: 326
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 293
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-523-419-293


  Query Match             100.0%;  Score 51;  DB 19;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.6e+06;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RMFPNAPYL 9
           |||||||||
Db       1 RMFPNAPYL 9

RESULT  13
US-09-625-963-1
; Sequence 1, Application US/09625963
; GENERAL INFORMATION:
;  APPLICANT:  Strauss, Hans Josef
;  APPLICANT:  Gao, Liquan
;  TITLE OF INVENTION:  Immunotherapeutic Methods Using Epitopes of WT-1 and
;  TITLE OF INVENTION:  GATA-1
;  FILE REFERENCE:  ICI 101
;  CURRENT APPLICATION NUMBER: US/09/625,963
;  CURRENT FILING DATE:  2000-07-26
;  PRIOR APPLICATION NUMBER:  PCT/GB99/03572
;  PRIOR FILING DATE:  1999-11-02
;  PRIOR APPLICATION NUMBER:  GB9823897.5
;  PRIOR FILING DATE:  1998-11-02
;  NUMBER OF SEQ ID NOS:  5
;  SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-625-963-1


  Query Match             100.0%;  Score 51;  DB 20;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.6e+06;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RMFPNAPYL 9
           |||||||||
Db       1 RMFPNAPYL 9

RESULT  14
US-09-679-339-185
; Sequence 185, Application US/09679339
; GENERAL INFORMATION:
;  APPLICANT:  Skeiky, Yasir
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Cheever, Martin
;  APPLICANT:  Reed, Steven G.
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR WT1
;  TITLE OF INVENTION:  SPECIFIC IMMUNOTHERAPY
;  FILE REFERENCE:  210121.533
;  CURRENT APPLICATION NUMBER: US/09/679,339
;  CURRENT FILING DATE:  2000-10-04
;  NUMBER OF SEQ ID NOS:  357
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 185
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Homo sapien
US-09-679-339-185

  Query Match             100.0%;  Score 51;  DB 20;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.6e+06;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RMFPNAPYL 9
           |||||||||
Db       1 RMFPNAPYL 9

RESULT  15
US-09-679-339-293
; Sequence 293, Application US/09679339
; GENERAL INFORMATION:
;  APPLICANT:  Skeiky, Yasir
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Cheever, Martin
;  APPLICANT:  Reed, Steven G.
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR WT1
;  TITLE OF INVENTION:  SPECIFIC IMMUNOTHERAPY
;  FILE REFERENCE:  210121.533
;  CURRENT APPLICATION NUMBER: US/09/679,339
;  CURRENT FILING DATE:  2000-10-04
;  NUMBER OF SEQ ID NOS:  357
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 293
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-679-339-293


  Query Match             100.0%;  Score 51;  DB 20;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.6e+06;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RMFPNAPYL 9
           |||||||||
Db       1 RMFPNAPYL 9


Search completed: October 23, 2001, 13:33:06
Job time: 283 sec
```

```
                      GenCore version 4.5
              Copyright (c) 1993 - 2000  Compugen Ltd.

OM protein - protein search, using sw model

Run on:         October 23, 2001, 13:22:12 ; Search time 31.82 Seconds
                                             (without alignments)
                                             17.147 Million cell updates/sec

Title:          US-09-625-963-1
Perfect score:  51
Sequence:       1 RMFPNAPYL 9

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched:       412676 seqs, 60623988 residues

Total number of hits satisfying chosen parameters:      75637

Minimum DB seq length: 0
Maximum DB seq length: 9

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       A_Geneseq_0601:*
               1:  /SIDS8/gcgdata/geneseq/geneseqp/AA1980.DAT:*
               2:  /SIDS8/gcgdata/geneseq/geneseqp/AA1981.DAT:*
               3:  /SIDS8/gcgdata/geneseq/geneseqp/AA1982.DAT:*
               4:  /SIDS8/gcgdata/geneseq/geneseqp/AA1983.DAT:*
               5:  /SIDS8/gcgdata/geneseq/geneseqp/AA1984.DAT:*
               6:  /SIDS8/gcgdata/geneseq/geneseqp/AA1985.DAT:*
               7:  /SIDS8/gcgdata/geneseq/geneseqp/AA1986.DAT:*
               8:  /SIDS8/gcgdata/geneseq/geneseqp/AA1987.DAT:*
               9:  /SIDS8/gcgdata/geneseq/geneseqp/AA1988.DAT:*
              10:  /SIDS8/gcgdata/geneseq/geneseqp/AA1989.DAT:*
              11:  /SIDS8/gcgdata/geneseq/geneseqp/AA1990.DAT:*
              12:  /SIDS8/gcgdata/geneseq/geneseqp/AA1991.DAT:*
              13:  /SIDS8/gcgdata/geneseq/geneseqp/AA1992.DAT:*
              14:  /SIDS8/gcgdata/geneseq/geneseqp/AA1993.DAT:*
              15:  /SIDS8/gcgdata/geneseq/geneseqp/AA1994.DAT:*
              16:  /SIDS8/gcgdata/geneseq/geneseqp/AA1995.DAT:*
              17:  /SIDS8/gcgdata/geneseq/geneseqp/AA1996.DAT:*
              18:  /SIDS8/gcgdata/geneseq/geneseqp/AA1997.DAT:*
              19:  /SIDS8/gcgdata/geneseq/geneseqp/AA1998.DAT:*
              20:  /SIDS8/gcgdata/geneseq/geneseqp/AA1999.DAT:*
              21:  /SIDS8/gcgdata/geneseq/geneseqp/AA2000.DAT:*
              22:  /SIDS8/gcgdata/geneseq/geneseqp/AA2001.DAT:*

    Pred. No. is the number of results predicted by chance to have a
    score greater than or equal to the score of the result being printed,
    and is derived by analysis of the total score distribution.

                               SUMMARIES
               %
Result        Query
  No.  Score  Match Length DB  ID                 Description
 ---------------------------------------------------------------------
   1     51  100.0      9  21  AAY94202           Human cytotoxic T
   2     51  100.0      9  21  AAY98670           WT1 derived immuno
   3     51  100.0      9  21  AAY98778           WT1 derived immuno
   4     51  100.0      9  21  AAY80200           Human Wilms' tumou
   5     47   92.2      9  21  AAY98523           WT1 derived immuno
   6     46   90.2      9  21  AAY98809           WT1 related peptid
   7     41   80.4      9  21  AAY98564           WT1 derived immuno
   8     41   80.4      9  21  AAY98752           WT1 derived immuno
   9     40   78.4      9  21  AAY98808           WT1 related peptid
  10     35   68.6      9  21  AAY98810           WT1 related peptid
  11     33   64.7      9  21  AAY98583           WT1 derived immuno
  12     33   64.7      9  21  AAY98754           WT1 derived immuno
  13     29   56.9      9  21  AAY98697           WT1 derived immuno
  14     28   54.9      9  18  AAW37017           TRP-2 derived pote
  15     28   54.9      9  19  AAW54658           Peptide from Measl
  16     28   54.9      9  20  AAY10403           T cell epitope/MHC
  17     28   54.9      9  21  AAY98629           WT1 derived immuno
  18     28   54.9      9  21  AAY98767           WT1 derived immuno
  19     27   52.9      8  16  AAR84531           Hepatitis B virus
  20     27   52.9      8  20  AAY48023           Immunogenic peptid
  21     26   51.0      6  19  AAW47460           Prothrombin/thromb
  22     26   51.0      9  20  AAY46620           Immunogenic peptid
  23     26   51.0      9  20  AAY46621           Immunogenic peptid
  24     26   51.0      9  20  AAY46625           Immunogenic peptid
  25     25   49.0      9  18  AAW37019           TRP-2 derived pote
  26     25   49.0      9  22  AAB76105           Tumour associated
  27     24   47.1      7  19  AAW75778           Methionyl oligopep
  28     24   47.1      9  15  AAY38094           Hepatitis B virus-
  29     24   47.1      9  15  AAR61680           HLA-A2.1 algorithm
  30     24   47.1      9  15  AAR73698           Antigen fragment 1
  31     24   47.1      9  16  AAR84772           MART-1 melanoma an
  32     24   47.1      9  18  AAW37015           TRP-2 derived pote
  33     24   47.1      9  20  AAY45662           Immunogenic peptid
  34     24   47.1      9  21  AAY73047           Hepatitis B virus
  35     23   45.1      7  11  AAR05940           Embryonic immunogl
  36     23   45.1      8  14  AAR35949           Hepatitis C virus
  37     23   45.1      8  14  AAR35950           Hepatitis C virus
  38     23   45.1      8  20  AAY13764           Human complement f
  39     23   45.1      8  22  AAB31526           Amino acid sequenc
  40     23   45.1      9  20  AAY46527           Immunogenic peptid
  41     23   45.1      9  20  AAY25555           Human MHC Class II
  42     23   45.1      9  20  AAY00821           HLA-A24 antigenic
  43     23   45.1      9  21  AAY98807           WT1 related peptid
  44     22   43.1      6  19  AAY21102           Human p53 cellular
  45     22   43.1      6  19  AAW47459           Prothrombin/thromb

                               ALIGNMENTS

RESULT   1
AAY94202
ID   AAY94202 standard; peptide; 9 AA.
XX
AC   AAY94202;
XX
DT   28-JUL-2000  (first entry)
XX
DE   Human cytotoxic T lymphocyte-recognised WT1 peptide WT126-34.
XX
KW   WT126-3; peptide; epitope; Wilm's tumour gene; leukaemia;
KW   breast cancer; melanoma; ovarian cancer; immunotherapy.
XX
OS   Homo sapiens.
XX
PN   WO200026249-A1.
XX
PD   11-MAY-2000.
XX
PF   02-NOV-1999;   99WO-GB03572.
XX
PR   02-NOV-1998;   98GB-0023897.
XX
PA   (IMCO-) IMPERIAL COLLEGE INNOVATIONS LTD.
XX
PI   Stauss HJ,  Gao L;
XX
DR   WPI; 2000-376123/32.
XX
PT   Novel peptides comprising WT-1 and GATA-1 epitopes, their fragments or
PT   variants, useful as vaccines for cancer immunotherapy  -
XX
PS   Claim 1; Page 74; 93pp; English.
XX
```

```
CC   The present sequence is peptide epitope WT126-34, produced by WT1
CC   expressing cells and found at residues 126-134 of the WT1 protein, which
CC   is recognised by cytotoxic T lymphocytes. WT1 is aberrantly expressed in
CC   leukaemias, breast cancer, melanoma and ovarian cancer. The peptide can
CC   be used as a vaccine to stimulate the elimination, by cytotoxic T
CC   lymphocytes, of cancer cells aberrantly expressing WT1. In addition, the
CC   nucleic acid encoding the peptide may also be used in the same manner.
CC   Alternatively, the peptide may be used in vitro to produce activated
CC   cytotoxic T lymphocytes.
XX
SQ   Sequence   9 AA;


  Query Match             100.0%;  Score 51;  DB 21;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 3.4e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RMFPNAPYL 9
              |||||||||
Db          1 rmfpnapyl 9


RESULT   2
AAY98670
ID   AAY98670 standard; Peptide; 9 AA.
XX
AC   AAY98670;
XX
DT   31-JUL-2000  (first entry)
XX
DE   WT1 derived immunogenic peptide SEQ ID NO:185.
XX
KW   WT1; immunotherapy; immunogenic; malignant disease; cancer; leukaemia;
KW   metastatic disease; mouse; human; Wilm's tumour; immune response;
KW   vaccine.
XX
OS   Homo sapiens.
XX
PN   WO200018795-A2.
XX
PD   06-APR-2000.
XX
PF   30-SEP-1999;   99WO-US22819.
XX
PR   30-SEP-1998;   98US-0164223.
PR   25-MAR-1999;   99US-0276484.
XX
PA   (CORI-) CORIXA CORP.
PA   (GAIG/) GAIGER A.
XX
PI   Gaiger A,  Cheever M;
XX
DR   WPI; 2000-293107/25.
XX
PT   Novel polypeptides comprising an immunogenic portion of a native WT1
PT   polypeptide, useful for inhibiting the development of malignant
PT   diseases associated with WT1 expression e.g. leukemia or cancer  -
XX
PS   Claim 4; Page 171; 193pp; English.
XX
CC   The present invention describes polypeptides (I) comprising an
CC   immunogenic portion of a native Wilm's Tumour gene product polypeptide,
CC   WT1, (or variants of the immunogenic portion retaining the ability to
CC   react with WT1-specific antisera and/or T-cell lines or clones) and
CC   comprising 16 consecutive amino acids (aa) or less of a native WT1
CC   polypeptide. The polypeptides are useful therapeutically and to
CC   manufacture medicaments for enhancing/inducing an immune response in
CC   patients. The polypeptides, mimetics or polynucleotides can be included
CC   with a carrier/excipient in pharmaceutical compositions or with a
CC   non-specific immune response enhancer (e.g. an adjuvant or enhancer
CC   preferentially enhancing a T cell response) in vaccines. Pharmaceutical
CC   compositions and vaccines can be administered to human patients to
CC   enhance or induce an immune response specific for WT1 or a cell
CC   expressing WT1, useful to inhibit the development of malignant diseases
CC   associated with WT1 expression, e.g. leukemia (especially acute/chronic
CC   myeloid leukemia or acute lymphocytic leukemia) or cancer (especially
CC   breast, lung, thyroid or gastrointestinal cancer, or a melanoma).
CC   AAY98501 to AAY98811 represent polypeptide sequences, and AAA13848 to
CC   AAA13862 represent PCR primers, used in the exemplification of the
CC   present invention.
XX
SQ   Sequence   9 AA;

  Query Match             100.0%;  Score 51;  DB 21;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 3.4e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RMFPNAPYL 9
              |||||||||
Db          1 rmfpnapyl 9


RESULT   3
AAY98778
ID   AAY98778 standard; Peptide; 9 AA.
XX
AC   AAY98778;
XX
DT   31-JUL-2000  (first entry)
XX
DE   WT1 derived immunogenic peptide SEQ ID NO:293.
XX
KW   WT1; immunotherapy; immunogenic; malignant disease; cancer; leukaemia;
KW   metastatic disease; mouse; human; Wilm's tumour; immune response;
KW   vaccine.
XX
OS   Mus musculus.
XX
PN   WO200018795-A2.
XX
PD   06-APR-2000.
XX
PF   30-SEP-1999;   99WO-US22819.
XX
PR   30-SEP-1998;   98US-0164223.
PR   25-MAR-1999;   99US-0276484.
XX
PA   (CORI-) CORIXA CORP.
PA   (GAIG/) GAIGER A.
XX
PI   Gaiger A,  Cheever M;
XX
DR   WPI; 2000-293107/25.
XX
PT   Novel polypeptides comprising an immunogenic portion of a native WT1
PT   polypeptide, useful for inhibiting the development of malignant
PT   diseases associated with WT1 expression e.g. leukemia or cancer  -
XX
PS   Claim 4; Page 186; 193pp; English.
XX
CC   The present invention describes polypeptides (I) comprising an
CC   immunogenic portion of a native Wilm's Tumour gene product polypeptide,
CC   WT1, (or variants of the immunogenic portion retaining the ability to
CC   react with WT1-specific antisera and/or T-cell lines or clones) and
CC   comprising 16 consecutive amino acids (aa) or less of a native WT1
CC   polypeptide. The polypeptides are useful therapeutically and to
CC   manufacture medicaments for enhancing/inducing an immune response in
CC   patients. The polypeptides, mimetics or polynucleotides can be included
CC   with a carrier/excipient in pharmaceutical compositions or with a
CC   non-specific immune response enhancer (e.g. an adjuvant or enhancer
CC   preferentially enhancing a T cell response) in vaccines. Pharmaceutical
CC   compositions and vaccines can be administered to human patients to
CC   enhance or induce an immune response specific for WT1 or a cell
CC   expressing WT1, useful to inhibit the development of malignant diseases
CC   associated with WT1 expression, e.g. leukemia (especially acute/chronic
```

```
CC   myeloid leukemia or acute lymphocytic leukemia) or cancer (especially
CC   breast, lung, thyroid or gastrointestinal cancer, or a melanoma).
CC   AAY98501 to AAY98811 represent polypeptide sequences, and AAA13848 to
CC   AAA13862 represent PCR primers, used in the exemplification of the
CC   present invention.
XX
SQ   Sequence   9 AA;


  Query Match             100.0%;  Score 51;  DB 21;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 3.4e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RMFPNAPYL 9
           |||||||||
Db       1 rmfpnapyl 9


RESULT   4
AAY80200
ID   AAY80200 standard; Peptide; 9 AA.
XX
AC   AAY80200;
XX
DT   24-MAY-2000  (first entry)
XX
DE   Human Wilms' tumour suppressor gene WT1 product peptide SEQ ID NO:5.
XX
KW   Wilm's tumour suppressor gene; WT1; cancer; antigen; vaccine; MHC;
KW   major histocompatibility complex; leukaemia; tumour; antitumour.
XX
OS   Homo-sapiens.
XX
PN   WO200006602-A1.
XX
PD   10-FEB-2000.
XX
PF   30-JUL-1999;   99WO-JP04130.
XX
PR   31-JUL-1998;   98JP-0218093.
XX
PA   (SUGI/) SUGIYAMA H.
XX
PI   Sugiyama H,  Oka Y;
XX
DR   WPI; 2000-195264/17.
XX
PT   Cancer antigens based on Wilm's tumor suppressor gene WT1 product or
PT   peptide derivatives, for cancer vaccines in treating leukemia and solid
PT   tumors e.g. stomach cancer, skin cancer, lung cancer and breast cancer
XX
PS   Claim 5; Page 18; 48pp; Japanese.
XX
CC   The present invention describes a cancer antigen containing the active
CC   component of Wilms' tumour suppressor gene WT1 product, or partial
CC   peptides, for cancer vaccines in treating leukaemia and solid tumours.
CC   The cancer antigens are useful for cancer vaccines in treating
CC   leukaemia, bone-marrow abnormal formation syndrome, malignant lymphoma,
CC   multiple myeloma, stomach cancer, cancer of the large intestine, lung
CC   cancer, breast cancer, blastoma, liver cancer, skin cancer, bladder
CC   cancer, prostate cancer, uterus cancer, cervical cancer, or ovary
CC   cancer. The present sequence represents a peptide from the human
CC   Wilms' tumour suppressor gene WT1 product.
XX
SQ   Sequence   9 AA;


  Query Match             100.0%;  Score 51;  DB 21;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 3.4e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RMFPNAPYL 9
           |||||||||
Db       1 rmfpnapyl 9


RESULT   5
AAY98523
ID   AAY98523 standard; Peptide; 9 AA.
XX
AC   AAY98523;
XX
DT   31-JUL-2000  (first entry)
XX
DE   WT1 derived immunogenic peptide SEQ ID NO:38.
XX
KW   WT1; immunotherapy; immunogenic; malignant disease; cancer; leukaemia;
KW   metastatic disease; mouse; human; Wilm's tumour; immune response;
KW   vaccine.
XX
OS   Homo sapiens.
XX
PN   WO200018795-A2.
XX
PD   06-APR-2000.
XX
PF   30-SEP-1999;   99WO-US22819.
XX
PR   30-SEP-1998;   98US-0164223.
PR   25-MAR-1999;   99US-0276484.
XX
PA   (CORI-) CORIXA CORP.
PA   (GAIG/) GAIGER A.
XX
PI   Gaiger A,  Cheever M;
XX
DR   WPI; 2000-293107/25.
XX
PT   Novel polypeptides comprising an immunogenic portion of a native WT1
PT   polypeptide, useful for inhibiting the development of malignant
PT   diseases associated with WT1 expression e.g. leukemia or cancer  -
XX
PS   Claim 4; Page 149; 193pp; English.
XX
CC   The present invention describes polypeptides (I) comprising an
CC   immunogenic portion of a native Wilm's Tumour gene product polypeptide,
CC   WT1, (or variants of the immunogenic portion retaining the ability to
CC   react with WT1-specific antisera and/or T-cell lines or clones) and
CC   comprising 16 consecutive amino acids (aa) or less of a native WT1
CC   polypeptide. The polypeptides are useful therapeutically and to
CC   manufacture medicaments for enhancing/inducing an immune response in
CC   patients. The polypeptides, mimetics or polynucleotides can be included
CC   with a carrier/excipient in pharmaceutical compositions or with a
CC   non-specific immune response enhancer (e.g. an adjuvant or enhancer
CC   preferentially enhancing a T cell response) in vaccines. Pharmaceutical
CC   compositions and vaccines can be administered to human patients to
CC   enhance or induce an immune response specific for WT1 or a cell
CC   expressing WT1, useful to inhibit the development of malignant diseases
CC   associated with WT1 expression, e.g. leukemia (especially acute/chronic
CC   myeloid leukemia or acute lymphocytic leukemia) or cancer (especially
CC   breast, lung, thyroid or gastrointestinal cancer, or a melanoma).
CC   AAY98501 to AAY98811 represent polypeptide sequences, and AAA13848 to
CC   AAA13862 represent PCR primers, used in the exemplification of the
CC   present invention.
XX
SQ   Sequence   9 AA;


  Query Match              92.2%;  Score 47;  DB 21;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 3.4e+05;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RMFPNAPY 8
           ||||||||
Db       2 rmfpnapy 9
```

```
RESULT   6
AAY98809
ID   AAY98809 standard; Peptide; 9 AA.
XX
AC   AAY98809;
XX
DT   31-JUL-2000  (first entry)
XX
DE   WT1 related peptide SEQ ID NO:324.
XX
KW   WT1; immunotherapy; immunogenic; malignant disease; cancer; leukaemia;
KW   metastatic disease; mouse; human; Wilm's tumour; immune response;
KW   vaccine.
XX
OS   Homo sapiens.
OS   Mus musculus.
XX
PN   WO200018795-A2.
XX
PD   06-APR-2000.
XX
PF   30-SEP-1999;   99WO-US22819.
XX
PR   30-SEP-1998;   98US-0164223.
PR   25-MAR-1999;   99US-0276484.
XX
PA   (CORI-) CORIXA CORP.
PA   (GAIG/) GAIGER A.
XX
PI   Gaiger A,  Cheever M;
XX
DR   WPI; 2000-293107/25.
XX
PT   Novel polypeptides comprising an immunogenic portion of a native WT1
PT   polypeptide, useful for inhibiting the development of malignant
PT   diseases associated with WT1 expression e.g. leukemia or cancer  -
XX
PS   Disclosure; Page 193; 193pp; English.
XX
CC   The present invention describes polypeptides (I) comprising an
CC   immunogenic portion of a native Wilm's Tumour gene product polypeptide,
CC   WT1, (or variants of the immunogenic portion retaining the ability to
CC   react with WT1-specific antisera and/or T-cell lines or clones) and
CC   comprising 16 consecutive amino acids (aa) or less of a native WT1
CC   polypeptide. The polypeptides are useful therapeutically and to
CC   manufacture medicaments for enhancing/inducing an immune response in
CC   patients. The polypeptides, mimetics or polynucleotides can be included
CC   with a carrier/excipient in pharmaceutical compositions or with a
CC   non-specific immune response enhancer (e.g. an adjuvant or enhancer
CC   preferentially enhancing a T cell response) in vaccines. Pharmaceutical
CC   compositions and vaccines can be administered to human patients to
CC   enhance or induce an immune response specific for WT1 or a cell
CC   expressing WT1, useful to inhibit the development of malignant diseases
CC   associated with WT1 expression, e.g. leukemia (especially acute/chronic
CC   myeloid leukemia or acute lymphocytic leukemia) or cancer (especially
CC   breast, lung, thyroid or gastrointestinal cancer, or a melanoma).
CC   AAY98501 to AAY98811 represent polypeptide sequences, and AAA13848 to
CC   AAA13862 represent PCR primers, used in the exemplification of the
CC   present invention.
XX
SQ   Sequence   9 AA;


  Query Match             90.2%;  Score 46;  DB 21;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 3.4e+05;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2 MFPNAPYL 9
           ||||||||
Db       1 mfpnapyl 8
```

```
RESULT   7
AAY98564
ID   AAY98564 standard; Peptide; 9 AA.
XX
AC   AAY98564;
XX
DT   31-JUL-2000  (first entry)
XX
DE   WT1 derived immunogenic peptide SEQ ID NO:79.
XX
KW   WT1; immunotherapy; immunogenic; malignant disease; cancer; leukaemia;
KW   metastatic disease; mouse; human; Wilm's tumour; immune response;
KW   vaccine.
XX
OS   Homo sapiens.
XX
PN   WO200018795-A2.
XX
PD   06-APR-2000.
XX
PF   30-SEP-1999;   99WO-US22819.
XX
PR   30-SEP-1998;   98US-0164223.
PR   25-MAR-1999;   99US-0276484.
XX
PA   (CORI-) CORIXA CORP.
PA   (GAIG/) GAIGER A.
XX
PI   Gaiger A,  Cheever M;
XX
DR   WPI; 2000-293107/25.
XX
PT   Novel polypeptides comprising an immunogenic portion of a native WT1
PT   polypeptide, useful for inhibiting the development of malignant
PT   diseases associated with WT1 expression e.g. leukemia or cancer  -
XX
PS   Claim 4; Page 155; 193pp; English.
XX
CC   The present invention describes polypeptides (I) comprising an
CC   immunogenic portion of a native Wilm's Tumour gene product polypeptide,
CC   WT1, (or variants of the immunogenic portion retaining the ability to
CC   react with WT1-specific antisera and/or T-cell lines or clones) and
CC   comprising 16 consecutive amino acids (aa) or less of a native WT1
CC   polypeptide. The polypeptides are useful therapeutically and to
CC   manufacture medicaments for enhancing/inducing an immune response in
CC   patients. The polypeptides, mimetics or polynucleotides can be included
CC   with a carrier/excipient in pharmaceutical compositions or with a
CC   non-specific immune response enhancer (e.g. an adjuvant or enhancer
CC   preferentially enhancing a T cell response) in vaccines. Pharmaceutical
CC   compositions and vaccines can be administered to human patients to
CC   enhance or induce an immune response specific for WT1 or a cell
CC   expressing WT1, useful to inhibit the development of malignant diseases
CC   associated with WT1 expression, e.g. leukemia (especially acute/chronic
CC   myeloid leukemia or acute lymphocytic leukemia) or cancer (especially
CC   breast, lung, thyroid or gastrointestinal cancer, or a melanoma).
CC   AAY98501 to AAY98811 represent polypeptide sequences, and AAA13848 to
CC   AAA13862 represent PCR primers, used in the exemplification of the
CC   present invention.
XX
SQ   Sequence   9 AA;


  Query Match             80.4%;  Score 41;  DB 21;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 3.4e+05;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3 FPNAPYL 9
           |||||||
Db       1 fpnapyl 7


RESULT   8
AAY98752
```

```
ID   AAY98752 standard; Peptide; 9 AA.
XX
AC   AAY98752;
XX
DT   31-JUL-2000  (first entry)
XX
DE   WT1 derived immunogenic peptide SEQ ID NO:267.
XX
KW   WT1; immunotherapy; immunogenic; malignant disease; cancer; leukaemia;
KW   metastatic disease; mouse; human; Wilm's tumour; immune response;
KW   vaccine.
XX
OS   Mus musculus.
XX
PN   WO200018795-A2.
XX
PD   06-APR-2000.
XX
PF   30-SEP-1999;   99WO-US22819.
XX
PR   30-SEP-1998;   98US-0164223.
PR   25-MAR-1999;   99US-0276484.
XX
PA   (CORI-) CORIXA CORP.
PA   (GAIG/) GAIGER A.
XX
PI   Gaiger A,  Cheever M;
XX
DR   WPI; 2000-293107/25.
XX
PT   Novel polypeptides comprising an immunogenic portion of a native WT1
PT   polypeptide, useful for inhibiting the development of malignant
PT   diseases associated with WT1 expression e.g. leukemia or cancer  -
XX
PS   Claim 4; Page 182; 193pp; English.
XX
CC   The present invention describes polypeptides (I) comprising an
CC   immunogenic portion of a native Wilm's Tumour gene product polypeptide,
CC   WT1, (or variants of the immunogenic portion retaining the ability to
CC   react with WT1-specific antisera and/or T-cell lines or clones) and
CC   comprising 16 consecutive amino acids (aa) or less of a native WT1
CC   polypeptide. The polypeptides are useful therapeutically and to
CC   manufacture medicaments for enhancing/inducing an immune response in
CC   patients. The polypeptides, mimetics or polynucleotides can be included
CC   with a carrier/excipient in pharmaceutical compositions or with a
CC   non-specific immune response enhancer (e.g. an adjuvant or enhancer
CC   preferentially enhancing a T cell response) in vaccines. Pharmaceutical
CC   compositions and vaccines can be administered to human patients to
CC   enhance or induce an immune response specific for WT1 or a cell
CC   expressing WT1, useful to inhibit the development of malignant diseases
CC   associated with WT1 expression, e.g. leukemia (especially acute/chronic
CC   myeloid leukemia or acute lymphocytic leukemia) or cancer (especially
CC   breast, lung, thyroid or gastrointestinal cancer, or a melanoma).
CC   AAY98501 to AAY98811 represent polypeptide sequences, and AAA13848 to
CC   AAA13862 represent PCR primers, used in the exemplification of the
CC   present invention.
XX
SQ   Sequence   9 AA;

  Query Match             80.4%;  Score 41;  DB 21;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 3.4e+05;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        3 FPNAPYL 9
            |||||||
Db        1 fpnapyl 7

RESULT   9
AAY98808
ID   AAY98808 standard; Peptide; 9 AA.
XX
AC   AAY98808;
XX
DT   31-JUL-2000  (first entry)
XX
DE   WT1 related peptide SEQ ID NO:323.
XX
KW   WT1; immunotherapy; immunogenic; malignant disease; cancer; leukaemia;
KW   metastatic disease; mouse; human; Wilm's tumour; immune response;
KW   vaccine.
XX
OS   Homo sapiens.
OS   Mus musculus.
XX
PN   WO200018795-A2.
XX
PD   06-APR-2000.
XX
PF   30-SEP-1999;   99WO-US22819.
XX
PR   30-SEP-1998;   98US-0164223.
PR   25-MAR-1999;   99US-0276484.
XX
PA   (CORI-) CORIXA CORP.
PA   (GAIG/) GAIGER A.
XX
PI   Gaiger A,  Cheever M;
XX
DR   WPI; 2000-293107/25.
XX
PT   Novel polypeptides comprising an immunogenic portion of a native WT1
PT   polypeptide, useful for inhibiting the development of malignant
PT   diseases associated with WT1 expression e.g. leukemia or cancer  -
XX
PS   Disclosure; Page 192; 193pp; English.
XX
CC   The present invention describes polypeptides (I) comprising an
CC   immunogenic portion of a native Wilm's Tumour gene product polypeptide,
CC   WT1, (or variants of the immunogenic portion retaining the ability to
CC   react with WT1-specific antisera and/or T-cell lines or clones) and
CC   comprising 16 consecutive amino acids (aa) or less of a native WT1
CC   polypeptide. The polypeptides are useful therapeutically and to
CC   manufacture medicaments for enhancing/inducing an immune response in
CC   patients. The polypeptides, mimetics or polynucleotides can be included
CC   with a carrier/excipient in pharmaceutical compositions or with a
CC   non-specific immune response enhancer (e.g. an adjuvant or enhancer
CC   preferentially enhancing a T cell response) in vaccines. Pharmaceutical
CC   compositions and vaccines can be administered to human patients to
CC   enhance or induce an immune response specific for WT1 or a cell
CC   expressing WT1, useful to inhibit the development of malignant diseases
CC   associated with WT1 expression, e.g. leukemia (especially acute/chronic
CC   myeloid leukemia or acute lymphocytic leukemia) or cancer (especially
CC   breast, lung, thyroid or gastrointestinal cancer, or a melanoma).
CC   AAY98501 to AAY98811 represent polypeptide sequences, and AAA13848 to
CC   AAA13862 represent PCR primers, used in the exemplification of the
CC   present invention.
XX
SQ   Sequence   9 AA;

  Query Match             78.4%;  Score 40;  DB 21;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 3.4e+05;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 RMFPNAP 7
            |||||||
Db        3 rmfpnap 9

RESULT  10
AAY98810
ID   AAY98810 standard; Peptide; 9 AA.
XX
AC   AAY98810;
```

```
XX
DT   31-JUL-2000  (first entry)
XX
DE   WT1 related peptide SEQ ID NO:325.
XX
KW   WT1; immunotherapy; immunogenic; malignant disease; cancer; leukaemia;
KW   metastatic disease; mouse; human; Wilm's tumour; immune response;
KW   vaccine.
XX
OS   Homo sapiens.
OS   Mus musculus.
XX
PN   WO200018795-A2.
XX
PD   06-APR-2000.
XX
PF   30-SEP-1999;   99WO-US22819.
XX
PR   30-SEP-1998;   98US-0164223.
PR   25-MAR-1999;   99US-0276484.
XX
PA   (CORI-) CORIXA CORP.
PA   (GAIG/) GAIGER A.
XX
PI   Gaiger A,  Cheever M;
XX
DR   WPI; 2000-293107/25.
XX
PT   Novel polypeptides comprising an immunogenic portion of a native WT1
PT   polypeptide, useful for inhibiting the development of malignant
PT   diseases associated with WT1 expression e.g. leukemia or cancer  -
XX
PS   Disclosure; Page 193; 193pp; English.
XX
CC   The present invention describes polypeptides (I) comprising an
CC   immunogenic portion of a native Wilm's Tumour gene product polypeptide,
CC   WT1, (or variants of the immunogenic portion retaining the ability to
CC   react with WT1-specific antisera and/or T-cell lines or clones) and
CC   comprising 16 consecutive amino acids (aa) or less of a native WT1
CC   polypeptide. The polypeptides are useful therapeutically and to
CC   manufacture medicaments for enhancing/inducing an immune response in
CC   patients. The polypeptides, mimetics or polynucleotides can be included
CC   with a carrier/excipient in pharmaceutical compositions or with a
CC   non-specific immune response enhancer (e.g. an adjuvant or enhancer
CC   preferentially enhancing a T cell response) in vaccines. Pharmaceutical
CC   compositions and vaccines can be administered to human patients to
CC   enhance or induce an immune response specific for WT1 or a cell
CC   expressing WT1, useful to inhibit the development of malignant diseases
CC   associated with WT1 expression, e.g. leukemia (especially acute/chronic
CC   myeloid leukemia or acute lymphocytic leukemia) or cancer (especially
CC   breast, lung, thyroid or gastrointestinal cancer, or a melanoma).
CC   AAY98501 to AAY98811 represent polypeptide sequences, and AAA13848 to
CC   AAA13862 represent PCR primers, used in the exemplification of the
CC   present invention.
XX
SQ   Sequence   9 AA;


  Query Match             68.6%;  Score 35;  DB 21;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 3.4e+05;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       4 PNAPYL 9
           ||||||
Db       1 pnapyl 6


RESULT 11
AAY98583
ID   AAY98583 standard; Peptide; 9 AA.
XX
AC   AAY98583;
XX
```

```
DT   31-JUL-2000  (first entry)
XX
DE   WT1 derived immunogenic peptide SEQ ID NO:98.
XX
KW   WT1; immunotherapy; immunogenic; malignant disease; cancer; leukaemia;
KW   metastatic disease; mouse; human; Wilm's tumour; immune response;
KW   vaccine.
XX
OS   Homo sapiens.
XX
PN   WO200018795-A2.
XX
PD   06-APR-2000.
XX
PF   30-SEP-1999;   99WO-US22819.
XX
PR   30-SEP-1998;   98US-0164223.
PR   25-MAR-1999;   99US-0276484.
XX
PA   (CORI-) CORIXA CORP.
PA   (GAIG/) GAIGER A.
XX
PI   Gaiger A,  Cheever M;
XX
DR   WPI; 2000-293107/25.
XX
PT   Novel polypeptides comprising an immunogenic portion of a native WT1
PT   polypeptide, useful for inhibiting the development of malignant
PT   diseases associated with WT1 expression e.g. leukemia or cancer  -
XX
PS   Claim 4; Page 158; 193pp; English.
XX
CC   The present invention describes polypeptides (I) comprising an
CC   immunogenic portion of a native Wilm's Tumour gene product polypeptide,
CC   WT1, (or variants of the immunogenic portion retaining the ability to
CC   react with WT1-specific antisera and/or T-cell lines or clones) and
CC   comprising 16 consecutive amino acids (aa) or less of a native WT1
CC   polypeptide. The polypeptides are useful therapeutically and to
CC   manufacture medicaments for enhancing/inducing an immune response in
CC   patients. The polypeptides, mimetics or polynucleotides can be included
CC   with a carrier/excipient in pharmaceutical compositions or with a
CC   non-specific immune response enhancer (e.g. an adjuvant or enhancer
CC   preferentially enhancing a T cell response) in vaccines. Pharmaceutical
CC   compositions and vaccines can be administered to human patients to
CC   enhance or induce an immune response specific for WT1 or a cell
CC   expressing WT1, useful to inhibit the development of malignant diseases
CC   associated with WT1 expression, e.g. leukemia (especially acute/chronic
CC   myeloid leukemia or acute lymphocytic leukemia) or cancer (especially
CC   breast, lung, thyroid or gastrointestinal cancer, or a melanoma).
CC   AAY98501 to AAY98811 represent polypeptide sequences, and AAA13848 to
CC   AAA13862 represent PCR primers, used in the exemplification of the
CC   present invention.
XX
SQ   Sequence   9 AA;

  Query Match             64.7%;  Score 33;  DB 21;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 3.4e+05;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RMFPNA 6
           ||||||
Db       4 rmfpna 9


RESULT 12
AAY98754
ID   AAY98754 standard; Peptide; 9 AA.
XX
AC   AAY98754;
XX
DT   31-JUL-2000  (first entry)
XX
```

```
DE   WT1 derived immunogenic peptide SEQ ID NO:269.
XX
KW   WT1; immunotherapy; immunogenic; malignant disease; cancer; leukaemia;
KW   metastatic disease; mouse; human; Wilm's tumour; immune response;
KW   vaccine.
XX
OS   Mus musculus.
XX
PN   WO200018795-A2.
XX
PD   06-APR-2000.
XX
PF   30-SEP-1999;   99WO-US22819.
XX
PR   30-SEP-1998;   98US-0164223.
PR   25-MAR-1999;   99US-0276484.
XX
PA   (CORI-) CORIXA CORP.
PA   (GAIG/) GAIGER A.
XX
PI   Gaiger A,  Cheever M;
XX
DR   WPI; 2000-293107/25.
XX
PT   Novel polypeptides comprising an immunogenic portion of a native WT1
PT   polypeptide, useful for inhibiting the development of malignant
PT   diseases associated with WT1 expression e.g. leukemia or cancer  -
XX
PS   Claim 4; Page 183; 193pp; English.
XX
CC   The present invention describes polypeptides (I) comprising an
CC   immunogenic portion of a native Wilm's Tumour gene product polypeptide,
CC   WT1, (or variants of the immunogenic portion retaining the ability to
CC   react with WT1-specific antisera and/or T-cell lines or clones) and
CC   comprising 16 consecutive amino acids (aa) or less of a native WT1
CC   polypeptide. The polypeptides are useful therapeutically and to
CC   manufacture medicaments for enhancing/inducing an immune response in
CC   patients. The polypeptides, mimetics or polynucleotides can be included
CC   with a carrier/excipient in pharmaceutical compositions or with a
CC   non-specific immune response enhancer (e.g. an adjuvant or enhancer
CC   preferentially enhancing a T cell response) in vaccines. Pharmaceutical
CC   compositions and vaccines can be administered to human patients to
CC   enhance or induce an immune response specific for WT1 or a cell
CC   expressing WT1, useful to inhibit the development of malignant diseases
CC   associated with WT1 expression, e.g. leukemia (especially acute/chronic
CC   myeloid leukemia or acute lymphocytic leukemia) or cancer (especially
CC   breast, lung, thyroid or gastrointestinal cancer, or a melanoma).
CC   AAY98501 to AAY98811 represent polypeptide sequences, and AAA13848 to
CC   AAA13862 represent PCR primers, used in the exemplification of the
CC   present invention.
XX
SQ   Sequence   9 AA;

  Query Match             64.7%;  Score 33;  DB 21;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 3.4e+05;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 RMFPNA 6
            ||||||
Db        4 rmfpna 9

RESULT  13
AAY98697
ID   AAY98697 standard; Peptide; 9 AA.
XX
AC   AAY98697;
XX
DT   31-JUL-2000  (first entry)
XX
DE   WT1 derived immunogenic peptide SEQ ID NO:212.
XX
KW   WT1; immunotherapy; immunogenic; malignant disease; cancer; leukaemia;
KW   metastatic disease; mouse; human; Wilm's tumour; immune response;
KW   vaccine.
XX
OS   Homo sapiens.
XX
PN   WO200018795-A2.
XX
PD   06-APR-2000.
XX
PF   30-SEP-1999;   99WO-US22819.
XX
PR   30-SEP-1998;   98US-0164223.
PR   25-MAR-1999;   99US-0276484.
XX
PA   (CORI-) CORIXA CORP.
PA   (GAIG/) GAIGER A.
XX
PI   Gaiger A,  Cheever M;
XX
DR   WPI; 2000-293107/25.
XX
PT   Novel polypeptides comprising an immunogenic portion of a native WT1
PT   polypeptide, useful for inhibiting the development of malignant
PT   diseases associated with WT1 expression e.g. leukemia or cancer  -
XX
PS   Claim 4; Page 174; 193pp; English.
XX
CC   The present invention describes polypeptides (I) comprising an
CC   immunogenic portion of a native Wilm's Tumour gene product polypeptide,
CC   WT1, (or variants of the immunogenic portion retaining the ability to
CC   react with WT1-specific antisera and/or T-cell lines or clones) and
CC   comprising 16 consecutive amino acids (aa) or less of a native WT1
CC   polypeptide. The polypeptides are useful therapeutically and to
CC   manufacture medicaments for enhancing/inducing an immune response in
CC   patients. The polypeptides, mimetics or polynucleotides can be included
CC   with a carrier/excipient in pharmaceutical compositions or with a
CC   non-specific immune response enhancer (e.g. an adjuvant or enhancer
CC   preferentially enhancing a T cell response) in vaccines. Pharmaceutical
CC   compositions and vaccines can be administered to human patients to
CC   enhance or induce an immune response specific for WT1 or a cell
CC   expressing WT1, useful to inhibit the development of malignant diseases
CC   associated with WT1 expression, e.g. leukemia (especially acute/chronic
CC   myeloid leukemia or acute lymphocytic leukemia) or cancer (especially
CC   breast, lung, thyroid or gastrointestinal cancer, or a melanoma).
CC   AAY98501 to AAY98811 represent polypeptide sequences, and AAA13848 to
CC   AAA13862 represent PCR primers, used in the exemplification of the
CC   present invention.
XX
SQ   Sequence   9 AA;

  Query Match             56.9%;  Score 29;  DB 21;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 3.4e+05;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 RMFPN 5
            |||||
Db        5 rmfpn 9

RESULT  14
AAW37017
ID   AAW37017 standard; Peptide; 9 AA.
XX
AC   AAW37017;
XX
DT   20-MAR-1998  (first entry)
XX
DE   TRP-2 derived potential cancer antigen 18, based on positions 197-205.
XX
KW   Tyrosinase related protein 2 gene; TRP-1; gp75; tumour antigen;
KW   tumour infiltrating lymphocyte; TIL; TIL586; cancer peptide; TRP-2;
```

```
KW   alternative reading frame; cancer detection; pre-cancer detection;
KW   melanoma.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 3
FT                   /label= G3F
FT                   /note= "wild type Gly199 substituted with Phe"
XX
PN   WO9729195-A2.
XX
PD   14-AUG-1997.
XX
PF   06-FEB-1997;   97WO-US02186.
XX
PR   04-OCT-1996;   96US-0725736.
PR   09-FEB-1996;   96US-0599602.
XX
PA   (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
PI   Rosenberg SA,  Wang R;
XX
DR   WPI; 1997-415349/38.
XX
PT   Cancer antigen peptide(s) derived from the tyrosinase-related
PT   protein 1 or 2  - useful for detecting, preventing or treating a
PT   cancer in a mammal, especially melanoma
XX
PS   Example 14; Page 56; 111pp; English.
XX
CC   Peptides AAW30829-38 and AAW37011-21 are modified versions of a peptide
CC   derived from positions 197-205 of the tyrosinase related protein 2
CC   (TRP-2). This region contains the peptide epitope of TRP-2 that is
CC   able to stimulate cytokine release by CTL cells. Apart from
CC   AAW30829-30 (these contain extra residues at the N-terminal (AAW30829)
CC   and the C-terminal (AAW30830)), the peptides were modified to contain
CC   substitutions at the anchor residues. Of all these peptides, only
CC   AAW30829-38 were able to stimulate cytokine release. Other antigenic
CC   peptides have also been identified from TRP-1. The nucleic acids
CC   encoding the cancer peptides or TRP-2 can be used to detect a cancer or
CC   pre-cancer in a mammal, especially by detecting the presence of the
CC   alterative ORF 3 of the TRP-1 gene or the sequence encoding the novel
CC   tumour antigen TRP-2. Vectors and recombinant viruses containing antigen
CC   peptide encoding nucleic acids, antibodies raised against the peptides,
CC   or the peptides themselves can be used to prevent or treat a cancer in a
CC   mammal, especially a melanoma.
XX
SQ   Sequence   9 AA;

  Query Match            54.9%;  Score 28;  DB 18;  Length 9;
  Best Local Similarity   57.1%;  Pred. No. 3.4e+05;
  Matches    4;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       2 MFPNAPY 8
             :||  ||
Db       2 lfpgrpy 8

RESULT  15
AAW54658
ID   AAW54658 standard; peptide; 9 AA.
XX
AC   AAW54658;
XX
DT   25-SEP-1998  (first entry)
XX
DE   Peptide from Measles F protein 438-446.
XX
KW   Mannose-antigen; antigen-presenting cell; mannosylated peptide; T cell;
KW   vaccine; treatment.
XX
OS   Synthetic.
XX
PN   WO9813378-A1.
XX
PD   02-APR-1998.
XX
PF   25-SEP-1997;   97WO-NL00536.
XX
PR   26-SEP-1996;   96EP-0202701.
XX
PA   (UYLE-) RIJKSUNIV LEIDEN.
XX
PI   Drijfhout JW,  Koning F;
XX
DR   WPI; 1998-230631/20.
XX
PT   Increasing uptake and presentation of antigen(s) - by adding mannose
PT   residue(s) to antigen for increasing T cell response, useful in,
PT   e.g. vaccines against viral infection(s)
XX
PS   Disclosure; Page 29; 47pp; English.
XX
CC   The peptides AAW54559-W54809 are examples of peptides to which at least
CC   1 (preferably 2) mannose can be attached to increase their uptake as
CC   antigens by antigen-presenting cells.  Uptake of agonist mannosylated
CC   peptides will increase the T cell response, whereas uptake of antagonist
CC   peptides blocks the T cell response.  Blocking binding of immunogenic
CC   autoantigens can be used in treatment of type I diabetes, rheumatoid
CC   arthritis, graft rejection etc., also to induce T-cell non-
CC   responsiveness.  Vaccines containing mannosylated antigen are used to
CC   prevent or treat infections by, e.g. bacteria, viruses, fungi, helminths
CC    and parasites.
XX
SQ   Sequence   9 AA;

  Query Match            54.9%;  Score 28;  DB 19;  Length 9;
  Best Local Similarity   55.6%;  Pred. No. 3.4e+05;
  Matches    5;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 RMFPNAPYL 9
             | :|:| ||
Db       1 rrypdavyl 9


Search completed: October 23, 2001, 13:28:56
Job time: 404 sec
```

```
                    GenCore version 4.5
           Copyright (c) 1993 - 2000  Compugen Ltd.


OM protein - protein search, using sw model

Run on:          October 23, 2001, 13:27:02 ; Search time 22.25 Seconds
                                              (without alignments)
                                              30.812 Million cell updates/sec

Title:           US-09-625-963-1
Perfect score:   51
Sequence:        1 RMFPNAPYL 9

Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched:        219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters:      787

Minimum DB seq length: 0
Maximum DB seq length: 9

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       PIR_68:*
                1:  pir1:*
                2:  pir2:*
                3:  pir3:*
                4:  pir4:*

     Pred. No. is the number of results predicted by chance to have a
     score greater than or equal to the score of the result being printed,
     and is derived by analysis of the total score distribution.

                              SUMMARIES
                 %
Result           Query
  No.    Score   Match Length DB  ID                   Description
----------------------------------------------------------------------
    1      19     37.3      9  2  S70345               amine oxidase (cop
    2      18     35.3      9  2  D48186               ATPase R1 subunit
    3      17     33.3      7  2  PH1408               Ig heavy chain V r
    4      17     33.3      9  2  JS0302               xenopsin-related p
    5      17     33.3      9  2  A60320               xenopsin-related p
    6      17     33.3      9  2  PL0139               carbon-monoxide de
    7      16     31.4      7  2  S21230               dermorphin (Trp-4,
    8      16     31.4      8  2  I57018               gene Cftr protein
    9      16     31.4      9  2  A44787               calliFMRFamide 10
   10      16     31.4      9  2  D41978               calliFMRFamide 4 -
   11      15     29.4      6  4  S15596               orf 3 rara 5'-regi
   12      15     29.4      9  2  S66607               quinoline 2-oxidor
   13      15     29.4      9  2  A61620               locustamyotropin I
   14      15     29.4      9  2  PH1591               Ig H chain V-D-J r
   15      14     27.5      5  2  JS0319               subesophageal gang
   16      14     27.5      8  2  S21288               lectin - potato (f
   17      14     27.5      8  2  PT0030               inulinase (EC 3.2.
   18      14     27.5      8  2  D47393               neuropeptide calla
   19      14     27.5      8  2  E47393               neuropeptide calla
   20      14     27.5      8  2  B45800               serum albumin - do
   21      14     27.5      8  4  I54017               granulocyte-colony
   22      14     27.5      9  2  S66419               tetrameric protein
   23      14     27.5      9  2  S70332               endosperm protein,
   24      14     27.5      9  2  A29477               diuretic neuropept
   25      14     27.5      9  2  S39766               cardioactive pepti
   26      14     27.5      9  2  S77984               cytochrome-c oxida
   27      14     27.5      9  2  S66635               alpha-2-macroglobu
   28      14     27.5      9  2  S10784               enamelin i - bovin
   29      13     25.5      4  2  S53508               starvation-induced
```

```
   30      13     25.5      6  2  A61049               halo-toxin - Pseud
   31      13     25.5      6  2  A44916               mosquitocidal toxi
   32      13     25.5      7  2  A15398               choline oxidase (E
   33      13     25.5      7  2  I50210               gene c-rel protein
   34      13     25.5      7  2  E48394               glycoprotein compo
   35      13     25.5      8  2  S08995               hypertrehalosemic
   36      13     25.5      8  2  S08996               hypertrehalosemic
   37      13     25.5      8  2  A49823               adipokinetic hormo
   38      13     25.5      8  2  B49823               adipokinetic hormo
   39      13     25.5      8  2  A44960               neuropeptide Led-C
   40      13     25.5      8  2  B44960               neuropeptide Led-C
   41      13     25.5      8  2  A43976               hypertrehalosemic
   42      13     25.5      8  2  B43976               hypertrehalosemic
   43      13     25.5      8  2  PH1407               Ig heavy chain V r
   44      13     25.5      8  2  PL0184               capsid protein VP-
   45      13     25.5      8  2  B39745               endoglycosylcerami


                             ALIGNMENTS


RESULT   1
S70345
amine oxidase (copper-containing) (EC 1.4.3.6) II - Aspergillus niger (fragments)
C;Species: Aspergillus niger
C;Date: 19-Mar-1998 #sequence_revision 17-Apr-1998 #text_change 17-Apr-1998
C;Accession: S70345
R;Frebort, I.; Pec, P.; Luhova, L.; Toyama, H.; Matsushita, K.; Hirota, S.; Kitagawa,
Biochim. Biophys. Acta 1295, 59-72, 1996
A;Title: Two amine oxidases from Aspergillus niger AKU 3302 contain topa quinone as t
A;Reference number: S70344; MUID:96283794
A;Accession: S70345
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-5;6-9 <FRE>
C;Keywords: oxidoreductase


  Query Match             37.3%;  Score 19;  DB 2;  Length 9;
  Best Local Similarity   60.0%;  Pred. No. 2.2e+05;
  Matches    3;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy        4 PNAPY 8
              || |
Db        5 PNXEY 9


RESULT   2
D48186
ATPase R1 subunit - wood tobacco (fragment)
C;Species: Nicotiana sylvestris (wood tobacco)
C;Date: 16-Feb-1994 #sequence_revision 18-Nov-1994 #text_change 23-Feb-1997
C;Accession: D48186
R;De Paepe, R.; Forchioni, A.; Chetrit, P.; Vedel, F.
Proc. Natl. Acad. Sci. U.S.A. 90, 5934-5938, 1993
A;Title: Specific mitochondrial proteins in pollen: presence of an additional ATP sy
A;Reference number: A48186; MUID:93317598
A;Accession: D48186
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-9 <DEl>
A;Experimental source: pollen
A;Note: sequence extracted from NCBI backbone (NCBIP:134871)


  Query Match             35.3%;  Score 18;  DB 2;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.2e+05;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        6 APY 8
            |||
Db        6 APY 8
```

```
RESULT   3
PH1408
Ig heavy chain V region - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 17-Mar-1999
C;Accession: PH1408; PH1405
R;Shirasawa, T.; Miyazoe, I.; Hagiwara, S.; Kimoto, H.; Shigemoto, K.; Taniguchi, M.; Ta
J. Exp. Med. 176, 1209-1214, 1992
A;Title: Heavy chain variable (VH) region diversity generated by VH gene replacement in
ia virus.
A;Reference number: PH1403; MUID:93018837
A;Accession: PH1408
A;Molecule type: DNA
A;Residues: 1-7 <SHI>
A;Experimental source: clone micro m+ 46-12-2
A;Accession: PH1405
A;Molecule type: DNA
A;Residues: 1-7 <SHI2>
A;Experimental source: clone micro m+ 46-6
C;Superfamily: immunoglobulin V region; immunoglobulin homology
C;Keywords: heterotetramer; immunoglobulin


  Query Match             33.3%;  Score 17;  DB 2;  Length 7;
  Best Local Similarity   75.0%;  Pred. No. 2.2e+05;
  Matches    3;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 RMFP 4
           | ||
Db       4 RRFP 7


RESULT   4
JS0302
xenopsin-related peptide 2 - turkey
N;Contains: xenopsin-related peptide 1
C;Species: Meleagris gallopavo (common turkey)
C;Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 18-Aug-2000
C;Accession: JS0302
R;Carraway, R.E.; Cochrane, D.E.; Mitra, S.P.
Regul. Pept. 22, 303-314, 1988
A;Title: Xenopsin-related peptide generated in avian gastric extracts.
A;Reference number: JS0302; MUID:89042995
A;Accession: JS0302
A;Molecule type: protein
A;Residues: 1-9 <CAR>
C;Comment: The peptides are present within several tissues primarily in large molecular
C;Superfamily: yeast coatomer complex alpha chain; WD repeat homology
C;Keywords: neuropeptide
F;1-9/Product: xenopsin-related peptide 2 #status experimental <XP2>
F;2-9/Product: xenopsin-related peptide 1 #status experimental <XP1>


  Query Match             33.3%;  Score 17;  DB 2;  Length 9;
  Best Local Similarity   33.3%;  Pred. No. 2.2e+05;
  Matches    2;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       4 PNAPYL 9
           |  |::
Db       3 PKRPWI 8


RESULT   5
A60320
xenopsin-related peptide 2 - rat
N;Contains: xenopsin-related peptide 1
C;Species: Rattus norvegicus (Norway rat)
C;Date: 10-Nov-1992 #sequence_revision 10-Nov-1992 #text_change 18-Aug-2000
C;Accession: A60320
R;Carraway, R.E.; Mitra, S.P.; Muraki, K.
Regul. Pept. 29, 229-239, 1990
A;Title: Isolation and structures of xenopsin-related peptides from rat stomach, liver a
```

```
A;Reference number: A60320; MUID:91018491
A;Accession: A60320
A;Molecule type: protein
A;Residues: 1-9 <CAR>
A;Note: the authors purified these peptides from pepsin-treated extracts of stomach,
C;Comment: Xenopsin and xenopsin-related peptides are similar to neurotensin in seque
C;Superfamily: yeast coatomer complex alpha chain; WD repeat homology
C;Keywords: neuropeptide
F;1-9/Product: xenopsin-related peptide 2 #status experimental <MAT1>
F;2-9/Product: xenopsin-related peptide 1 #status experimental <MAT2>

  Query Match             33.3%;  Score 17;  DB 2;  Length 9;
  Best Local Similarity   33.3%;  Pred. No. 2.2e+05;
  Matches    2;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       4 PNAPYL 9
           |  |::
Db       3 PKRPWI 8

RESULT   6
PL0139
carbon-monoxide dehydrogenase (EC 1.2.99.2) large chain - Pseudomonas carboxydoflava
C;Species: Pseudomonas carboxydoflava
C;Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 28-Apr-1993
C;Accession: PL0139
R;Kraut, M.; Hugendieck, I.; Herwig, S.; Meyer, O.
Arch. Microbiol. 152, 335-341, 1989
A;Title: Homology and distribution of CO dehydrogenase structural genes in carboxydot
A;Reference number: PL0138; MUID:90055678
A;Accession: PL0139
A;Molecule type: protein
A;Residues: 1-9 <KRA>
A;Note: 2-Met is also found
C;Comment: Cardon-monoxide dehydrogenase consists of three polypeptide chains: large,
C;Keywords: oxidoreductase


  Query Match             33.3%;  Score 17;  DB 2;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.2e+05;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       5 NAP 7
           |||
Db       2 NAP 4


RESULT   7
S21230
dermorphin (Trp-4, Asn-7) [validated] - two-colored leaf frog (fragment)
C;Species: Phyllomedusa bicolor (two-colored leaf frog)
C;Date: 19-Mar-1997 #sequence_revision 10-Oct-1997 #text_change 18-Aug-2000
C;Accession: S21230
R;Mignogna, G.; Severini, C.; Simmaco, M.; Negri, L.; Falconieri Erspamer, G.; Kreil,
FEBS Lett. 302, 151-154, 1992
A;Title: Identification and characterization of two dermorphins from skin extracts of
A;Reference number: S21152; MUID:92339502
A;Accession: S21230
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-7 <MIG>
C;Superfamily: dermorphin precursor; dermorphin precursor amino-terminal homology


  Query Match             31.4%;  Score 16;  DB 2;  Length 7;
  Best Local Similarity   66.7%;  Pred. No. 2.2e+05;
  Matches    2;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3 FPN 5
           :||
Db       5 YPN 7
```

```
RESULT   8
I57018
gene Cftr protein - mouse (fragment)
C;Species: Mus sp. (mouse)
C;Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 28-Feb-1997
C;Accession: I57018
R;Dorin, J.R.; Stevenson, B.J.; Fleming, S.; Alton, E.W.; Dickinson, P.; Porteous, D.J.
Mamm. Genome 5, 465-472, 1994
A;Title: Long-term survival of the exon 10 insertional cystic fibrosis mutant mouse is a
A;Reference number: I57018; MUID:95037043
A;Accession: I57018
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-8 <RES>
A;Cross-references: GB:S74246; NID:g710482
C;Genetics:
A;Gene: Cftr


  Query Match             31.4%;  Score 16;  DB 2;  Length 8;
  Best Local Similarity   50.0%;  Pred. No. 2.2e+05;
  Matches    2;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       4 PNAP 7
           |::|
Db       2 PDSP 5


RESULT   9
A44787
calliFMRFamide 10 - bluebottle fly (Calliphora vomitoria)
C;Species: Calliphora vomitoria
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 17-Mar-1999
C;Accession: A44787
R;Duve, H.; Johnsen, A.H.; Sewell, J.C.; Scott, A.G.; Orchard, I.; Rehfeld, J.F.; Thorpe
Proc. Natl. Acad. Sci. U.S.A. 89, 2326-2330, 1992
A;Title: Isolation, structure, and activity of -Phe-Met-Arg-Phe-NH-2 neuropeptides (desi
A;Reference number: A41978; MUID:92196111
A;Accession: A44787
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-9 <DUV>
C;Keywords: amidated carboxyl end; neuropeptide
F;9/Modified site: amidated carboxyl end (Phe) #status experimental


  Query Match             31.4%;  Score 16;  DB 2;  Length 9;
  Best Local Similarity   33.3%;  Pred. No. 2.2e+05;
  Matches    2;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       4 PNAPYL 9
           ||  ::
Db       2 PNRDFM 7


RESULT  10
D41978
calliFMRFamide 4 - bluebottle fly (Calliphora vomitoria)
C;Species: Calliphora vomitoria
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 17-Mar-1999
C;Accession: D41978
R;Duve, H.; Johnsen, A.H.; Sewell, J.C.; Scott, A.G.; Orchard, I.; Rehfeld, J.F.; Thorpe
Proc. Natl. Acad. Sci. U.S.A. 89, 2326-2330, 1992
A;Title: Isolation, structure, and activity of -Phe-Met-Arg-Phe-NH-2 neuropeptides (desi
A;Reference number: A41978; MUID:92196111
A;Accession: D41978
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-9 <DUV>
C;Keywords: amidated carboxyl end; neuropeptide
F;9/Modified site: amidated carboxyl end (Phe) #status experimental
```

```
  Query Match             31.4%;  Score 16;  DB 2;  Length 9;
  Best Local Similarity   33.3%;  Pred. No. 2.2e+05;
  Matches    2;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       4 PNAPYL 9
           ||  ::
Db       2 PNQDFM 7


RESULT  11
S15596
orf 3 rara 5'-region - human
C;Species: Homo sapiens (man)
C;Date: 04-Jun-1999 #sequence_revision 04-Jun-1999 #text_change 28-Jun-1999
C;Accession: S15596
R;Brand, N.J.; Petkovich, M.; Chambon, P.
Nucleic Acids Res. 18, 6799-6806, 1990
A;Title: Characterization of a functional promoter for the human retinoic acid recept
A;Reference number: S15594; MUID:91088249
A;Accession: S15596
A;Molecule type: DNA
A;Residues: 1-6 <BRA>
A;Cross-references: EMBL:X56058; NID:g35876
A;Note: this ORF from Fig. 2 is not annotated in GenBank entry HSRARA2, release 111.0
C;Comment: This sequence is not thought to be translated.
C;Genetics:
A;Gene: GDB:RARA
A;Cross-references: GDB:120337; OMIM:180240
A;Map position: 17q12-17q12


  Query Match             29.4%;  Score 15;  DB 4;  Length 6;
  Best Local Similarity   60.0%;  Pred. No. 2.2e+05;
  Matches    3;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 MFPNA 6
           | |:|
Db       1 MAPSA 5


RESULT  12
S66607
quinoline 2-oxidoreductase beta chain - Comamonas testosteroni (fragment)
C;Species: Comamonas testosteroni
C;Date: 15-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 17-Mar-1999
C;Accession: S66607
R;Schach, S.; Tshisuaka, B.; Fetzner, S.; Lingens, F.
Eur. J. Biochem. 232, 536-544, 1995
A;Title: Quinoline 2-oxidoreductase and 2-oxo-1,2-dihydroquinoline 5,6-dioxygenase fr
A;Reference number: S66606; MUID:96035889
A;Accession: S66607
A;Molecule type: protein
A;Residues: 1-9 <SCH>
A;Experimental source: strain 63


  Query Match             29.4%;  Score 15;  DB 2;  Length 9;
  Best Local Similarity   50.0%;  Pred. No. 2.2e+05;
  Matches    3;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       3 FPNAPY 8
           ||   |
Db       3 FPAFAY 8


RESULT  13
A61620
locustamyotropin III - migratory locust
C;Species: Locusta migratoria (migratory locust)
C;Date: 21-Jul-1995 #sequence_revision 28-Jul-1995 #text_change 11-Jul-1997
C;Accession: A61620
R;Schoofs, L.; Holman, G.M.; Hayes, T.K.; Nachman, R.J.; Kochansky, J.P.; De Loof, A.
```

```
Insect Biochem. Mol. Biol. 22, 447-452, 1992
A;Title: Isolation, identification and synthesis of locustamyotropin III and IV, two add
A;Reference number: A61620
A;Accession: A61620
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-9 <SCH>
C;Keywords: amidated carboxyl end; neuropeptide
F;9/Modified site: amidated carboxyl end (Leu) #status experimental


  Query Match             29.4%;  Score 15;  DB 2;  Length 9;
  Best Local Similarity   44.4%;  Pred. No. 2.2e+05;
  Matches    4;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy        1 RMFPNAPYL 9
            |  |  | |
Db        1 RQQPFVPRL 9


RESULT  14
PH1591
Ig H chain V-D-J region (wild-type clone 142) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 02-Jun-1994 #sequence_revision 02-Jun-1994 #text_change 17-Mar-1999
C;Accession: PH1591
R;Levinson, D.A.; Campos-Torres, J.; Leder, P.
J. Exp. Med. 178, 317-329, 1993
A;Title: Molecular characterization of transgene-induced immunodeficiency in B-less mice
A;Reference number: PH1580; MUID:93301609
A;Accession: PH1591
A;Molecule type: DNA
A;Residues: 1-9 <LEV>
A;Experimental source: bone marrow pre-B lymphocyte
C;Keywords: immunoglobulin


  Query Match             29.4%;  Score 15;  DB 2;  Length 9;
  Best Local Similarity   66.7%;  Pred. No. 2.2e+05;
  Matches    2;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        6 APY 8
            :||
Db        5 SPY 7


RESULT  15
JS0319
subesophageal ganglion pentapeptide - house cricket
C;Species: Acheta domesticus (house cricket)
C;Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 20-Jun-2000
C;Accession: JS0319
R;Wicker, C.; Wicker, C.
Comp. Biochem. Physiol. C 88, 185-187, 1987
A;Title: Isolation and structure of a peptide isolated from the suboesophageal ganglion
A;Reference number: JS0319
A;Accession: JS0319
A;Molecule type: protein
A;Residues: 1-5 <WIC>


  Query Match             27.5%;  Score 14;  DB 2;  Length 5;
  Best Local Similarity   66.7%;  Pred. No. 2.2e+05;
  Matches    2;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        6 APY 8
            ||:
Db        3 APF 5


Search completed: October 23, 2001, 13:29:50
Job time: 168 sec
```

```
                       GenCore version 4.5
               Copyright (c) 1993 - 2000  Compugen Ltd.


OM protein - protein search, using sw model

Run on:          October 23, 2001, 13:29:28 ; Search time 12.6 Seconds
                                              (without alignments)
                                              24.468 Million cell updates/sec

Title:           US-09-625-963-1
Perfect score:   51
Sequence:        1 RMFPNAPYL 9

Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched:        93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters:      213

Minimum DB seq length: 0
Maximum DB seq length: 9

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       SwissProt_39:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                                SUMMARIES
                   %
Result           Query
  No.    Score   Match  Length DB  ID                    Description
----------------------------------------------------------------------------
    1       22    43.1      9  1  UPA7_HUMAN            P30093 homo sapien
    2       18    35.3      7  1  GFRP_MOUSE            P99025 mus musculu
    3       18    35.3      7  1  UN06_PINPS            P81675 pinus pinas
    4       18    35.3      8  1  B44K_PORGI            P81886 porphyromon
    5       17    33.3      9  1  DCML_PSECF            P19913 pseudomonas
    6       17    33.3      9  1  OXYF_SCYCA            P42997 scyliorhinu
    7       17    33.3      9  1  OXYT_OCTVU            P80027 octopus vul
    8       16    31.4      9  1  FAR4_CALVO            P41859 calliphora
    9       16    31.4      9  1  FARA_CALVO            P41865 calliphora
   10       16    31.4      9  1  FLA2_TREHY            P80159 treponema h
   11       16    31.4      9  1  UPA3_HUMAN            P30089 homo sapien
   12       15    29.4      8  1  ALL7_CARMA            P81809 carcinus ma
   13       15    29.4      9  1  LMT3_LOCMI            P41489 locusta mig
   14       14    27.5      5  1  SUGA_ACHDO            P19991 acheta dome
   15       14    27.5      7  1  ALL3_CARMA            P81806 carcinus ma
   16       14    27.5      7  1  ALL4_CARMA            P81807 carcinus ma
   17       14    27.5      7  1  ALL5_CARMA            P81808 carcinus ma
   18       14    27.5      7  1  FAR1_HELTI            P41871 helisoma tr
   19       14    27.5      8  1  AL15_CARMA            P81818 carcinus ma
   20       14    27.5      8  1  AL16_CARMA            P81819 carcinus ma
   21       14    27.5      8  1  ALL3_CYDPO            P82154 cydia pomon
   22       14    27.5      8  1  ALL4_CALVO            P41840 calliphora
   23       14    27.5      8  1  ALL4_CYDPO            P82155 cydia pomon
   24       14    27.5      8  1  ALL5_CALVO            P41841 calliphora
   25       14    27.5      8  1  ALL8_CARMA            P81811 carcinus ma
   26       14    27.5      8  1  ALL9_CARMA            P81812 carcinus ma
   27       14    27.5      8  1  FUSS_FUSSO            P81010 fusarium so
   28       14    27.5      9  1  AL10_CARMA            P81813 carcinus ma
   29       14    27.5      9  1  COXE_THUOB            P80975 thunnus obe
   30       14    27.5      9  1  DNF1_LOCMI            P16339 locusta mig
   31       14    27.5      9  1  NEUX_HUMAN            P04277 homo sapien
   32       14    27.5      9  1  NEUX_RAT              P11382 rattus norv
   33       14    27.5      9  1  TKC1_CALVO            P41517 calliphora
   34       13    25.5      5  1  PAP2_PARMA            P81864 pardachirus
   35       13    25.5      7  1  CHOX_ALCSP            P16101 alcaligenes
   36       13    25.5      7  1  FAR4_PANRE            P41875 panagrellus
   37       13    25.5      7  1  MNP1_LEPDE            P42984 leptinotars
   38       13    25.5      7  1  UF04_MOUSE            P38642 mus musculu
   39       13    25.5      8  1  FAR1_PANRE            P41872 panagrellus
   40       13    25.5      8  1  HTF1_PERAM            P04548 periplaneta
   41       13    25.5      8  1  HTF2_PERAM            P04549 periplaneta
   42       13    25.5      8  1  HTF_TENMO             P25419 tenebrio mo
   43       13    25.5      8  1  UF06_MOUSE            P38644 mus musculu
   44       13    25.5      8  1  UPA1_HUMAN            P30087 homo sapien
   45       13    25.5      9  1  CCAP_CARMA            P38556 carcinus ma


                              ALIGNMENTS


RESULT   1
UPA7_HUMAN
ID   UPA7_HUMAN     STANDARD;      PRT;     9 AA.
AC   P30093;
DT   01-APR-1993 (Rel. 25, Created)
DT   01-APR-1993 (Rel. 25, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   UNKNOWN PROTEIN FROM 2D-PAGE OF PLASMA (SPOT 18) (FRAGMENT).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Plasma;
RX   MEDLINE=93092937; PubMed=1459097;
RA   Hughes G.J., Frutiger S., Paquet N., Ravier F., Pasquali C.,
RA   Sanchez J.-C., James R., Tissot J.-D., Bjellqvist B.,
RA   Hochstrasser D.F.;
RT   "Plasma protein map: an update by microsequencing.";
RL   Electrophoresis 13:707-714(1992).
CC   -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN
CC       PROTEIN IS: 5.05, ITS MW IS: 37 KDA.
DR   SWISS-2DPAGE; P30093; HUMAN.
FT   NON_TER       1      1
FT   UNSURE        5      5
FT   NON_TER       9      9
SQ   SEQUENCE   9 AA;  1042 MW;  5C14477AEB0772C7 CRC64;


  Query Match             43.1%;  Score 22;  DB 1;  Length 9;
  Best Local Similarity   42.9%;  Pred. No. 9.3e+04;
  Matches    3;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          2 MFPNAPY 8
              : |  ||
Db          2 LVPEXPY 8


RESULT   2
GFRP_MOUSE
ID   GFRP_MOUSE     STANDARD;      PRT;     7 AA.
AC   P99025;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   GTP CYCLOHYDROLASE I FEEDBACK REGULATORY PROTEIN (P35) (FRAGMENT).
GN   GCHFR OR GFRP.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Liver;
RA   Sanchez J.-C., Rouge V., Frutiger S., Hughes G., Yan J.X.,
```

```
RA   Hoogland C., Appel R.D., Binz P.-A., Hochstrasser D.F.,
RA   Cowthorne M.;
RL   Submitted (AUG-1998) to the SWISS-PROT data bank.
CC   -!- FUNCTION: MEDIATES TETRAHYDROBIOPTERIN INHIBITION OF GTP
CC       CYCLOHYDROLASE I. THIS INHIBITION IS REVERSED BY L-PHENYLALANINE
CC       (BY SIMILARITY).
CC   -!- SUBUNIT: HOMODIMER (BY SIMILARITY).
DR   SWISS-2DPAGE; P99025; MOUSE.
FT   INIT_MET      0      0
FT   NON_TER       7      7
SQ   SEQUENCE   7 AA;  806 MW;  71B5B057273B4700 CRC64;

  Query Match             35.3%;  Score 18;  DB 1;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 9.3e+04;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       7 PYL 9
           |||
Db       1 PYL 3


RESULT   3
UN06_PINPS
ID   UN06_PINPS     STANDARD;      PRT;     7 AA.
AC   P81675;
DT   15-JUL-1999 (Rel. 38, Created)
DT   15-JUL-1999 (Rel. 38, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   UNKNOWN PROTEIN FROM 2D-PAGE OF NEEDLES (N141) (FRAGMENT).
OS   Pinus pinaster (Maritime pine).
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Coniferopsida; Coniferales; Pinaceae; Pinus.
OX   NCBI_TaxID=71647;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Needle;
RX   MEDLINE=99274088; PubMed=10344291;
RA   Costa P., Pionneau C., Bauw G., Dubos C., Bahrman N., Kremer A.,
RA   Frigerio J.-M., Plomion C.;
RT   "Separation and characterization of needle and xylem maritime pine
RT   proteins.";
RL   Electrophoresis 20:1098-1108(1999).
CC   -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN
CC       PROTEIN IS: 6.6, ITS MW IS: 25 KDA.
FT   NON_TER       1      1
FT   NON_TER       7      7
SQ   SEQUENCE   7 AA;  823 MW;  69D76724486B5740 CRC64;

  Query Match             35.3%;  Score 18;  DB 1;. Length 7;
  Best Local Similarity   28.6%;  Pred. No. 9.3e+04;
  Matches    2;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       2 MFPNAPY 8
           :: | |:
Db       1 LYGNLPF 7


RESULT   4
B44K_PORGI
ID   B44K_PORGI     STANDARD;      PRT;     8 AA.
AC   P81886;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   44 KDA IMMUNOGENIC PROTEIN (FRAGMENT).
OS   Porphyromonas gingivalis (Bacteroides gingivalis).
OC   Bacteria; CFB group; Bacteroidaceae; Porphyromonas.
OX   NCBI_TaxID=837;
RN   [1]
RP   SEQUENCE.
RC   STRAIN=VPB_3492;
RX   MEDLINE=20198497; PubMed=10731616;
RA   Norris J.M., Love D.N.;
RT   "Serum antibody responses of cats to soluble whole cell antigens of
RT   feline porphyromonas gingivalis.";
RL   Vet. Microbiol. 73:37-49(2000).
CC   -!- SIMILARITY: TO P.GINGIVALIS HEMAGGLUTININ A.
KW   Antigen.
FT   NON_TER       8      8
SQ   SEQUENCE   8 AA;  989 MW;  9554540326CB476D CRC64;

  Query Match             35.3%;  Score 18;  DB 1;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 9.3e+04;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       6 APY 8
           |||
Db       1 APY 3


RESULT   5
DCML_PSECF
ID   DCML_PSECF     STANDARD;      PRT;     9 AA.
AC   P19913;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   01-DEC-1992 (Rel. 24, Last annotation update)
DE   CARBON MONOXIDE DEHYDROGENASE LARGE CHAIN (EC 1.2.99.2) (FRAGMENT).
OS   Pseudomonas carboxydoflava.
OC   Bacteria; Proteobacteria; beta subdivision; Comamonadaceae;
OC   Hydrogenophaga.
OX   NCBI_TaxID=47421;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=90055678; PubMed=2818128;
RA   Kraut M., Hugendieck I., Herwig S., Meyer O.;
RT   "Homology and distribution of CO dehydrogenase structural genes in
RT   carboxydotrophic bacteria.";
RL   Arch. Microbiol. 152:335-341(1989).
CC   -!- CATALYTIC ACTIVITY: CO + H(2)O + ACCEPTOR = CO(2) + REDUCED
CC       ACCEPTOR.
CC   -!- COFACTOR: MOLYBDENUM.
CC   -!- SUBUNIT: CONSISTS OF THREE POLYPEPTIDE CHAINS: LARGE, MEDIUM, AND
CC       SMALL.
DR   PIR; PL0139; PL0139.
KW   Oxidoreductase; Molybdenum.
FT   VARIANT       2      2       N -> M.
FT   NON_TER       9      9
SQ   SEQUENCE   9 AA;  974 MW;  0224DAB6C2D76DD4 CRC64;

  Query Match             33.3%;  Score 17;  DB 1;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 9.3e+04;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       5 NAP 7
           |||
Db       2 NAP 4


RESULT   6
OXYF_SCYCA
ID   OXYF_SCYCA     STANDARD;      PRT;     9 AA.
AC   P42997;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   01-NOV-1995 (Rel. 32, Last annotation update)
DE   PHASVATOCIN.
OS   Scyliorhinus canicula (Spotted dogfish) (Spotted catshark).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC   Elasmobranchii; Galeomorphii; Galeoidea; Carcharhiniformes;
OC   Scyliorhinidae; Scyliorhinus.
OX   NCBI_TaxID=7830;
```

```
                                         .-T., Acher R.;
                                         es in cartilaginous
                                         ocin-like peptides
                                         hinus caniculus).";
RN   [1]                                 270(1994).
RP   SEQUENCE.
RC   TISSUE=Pituitary;                   Y ON RAT UTERUS.
RX   MEDLINE=95062247; PubMed=7972045;   SSIN/OXYTOCIN FAMILY.
RA   Chauvet J., Rouille Y., Chauve
RT   "Special evolution of neuroh
RT   fishes: asvatocin and pha           RM; 1.
RT   isolated from the spott
RL   Proc. Natl. Acad. Sc
CC   -!- FUNCTION: DIS                   AMIDATION.
CC   -!- SIMILARITY:                     17EDD76EB44449DB CRC64;
DR   InterPro; IP
DR   Pfam; PF00                 3%;  Score 17;  DB 1;  Length 9;
DR   PROSIT                     0%;  Pred. No. 9.3e+04;
KW   Horm              ive     0;  Mismatches    2;  Indels    0;  Gaps    0;
FT   DI
FT
SQ
Qy
Db

AC                0      STANDARD;        PRT;     9 AA.
DT   01-MAR-1992 (Rel. 21, Created)
DT   01-MAR-1992 (Rel. 21, Last sequence update)
DT   01-NOV-1995 (Rel. 32, Last annotation update)
DE   CEPHALOTOCIN.
OS   Octopus vulgaris (Octopus).
OC   Eukaryota; Metazoa; Mollusca; Cephalopoda; Coleoidea; Octopoda;
OC   Incirrata; Octopodidae; Octopus.
OX   NCBI_TaxID=6645;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Nerve endings;
RX   MEDLINE=92270139; PubMed=1589145;
RA   Reich G.;
RT   "A new peptide of the oxytocin/vasopressin family isolated from
RT   nerves of the cephalopod Octopus vulgaris.";
RL   Neurosci. Lett. 134:191-194(1992).
CC   -!- FUNCTION: HAS A ROLE IN THE NEUROSECRETORY SYSTEM OF THE VENA
CC       CAVA.
CC   -!- SIMILARITY: BELONGS TO THE VASOPRESSIN/OXYTOCIN FAMILY.
DR   InterPro; IPR000981; -.
DR   Pfam; PF00220; hormone4; 1.
DR   PROSITE; PS00264; NEUROHYPOPHYS_HORM; 1.
KW   Hormone; Amidation.
FT   DISULFID      1      6
FT   MOD_RES       9      9       AMIDATION.
SQ   SEQUENCE   9 AA;  1072 MW;  17FF476EB45409DB CRC64;

  Query Match             33.3%;  Score 17;  DB 1;  Length 9;
  Best Local Similarity   60.0%;  Pred. No. 9.3e+04;
  Matches    3;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

y          3 FPNAP 7
             | | |
b          3 FRNCP 7

RESULT 8
FAR4_CALVO
ID   FAR4_CALVO      STANDARD;      PRT;     9 AA.
```

```
AC   P41859;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   01-NOV-1995 (Rel. 32, Last annotation update)
DE   CALLIFMRFAMIDE 4.
OS   Calliphora vomitoria (Blue blowfly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta,
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha
OC   Oestroidea; Calliphoridae; Calliphora.
OX   NCBI_TaxID=27454;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Thoracic ganglion;
RX   MEDLINE=92196111; PubMed=1549595;
RA   Duve H., Johnsen A.H., Sewell J.C., Scott A.G., Orchard I.,
RA   Rehfeld J.F., Thorpe A.;
RT   "Isolation, structure, and activity of -Phe-Met-Arg-Phe-NH2
RT   neuropeptides (designated calliFMRFamides) from the blowfly
RT   Calliphora vomitoria.";
RL   Proc. Natl. Acad. Sci. U.S.A. 89:2326-2330(1992).
CC   -!- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)
CC       FAMILY.
DR   PIR; D41978; D41978.
KW   Neuropeptide; Amidation.
FT   MOD_RES       9      9       AMIDATION.
SQ   SEQUENCE   9 AA;  1182 MW;  31730699CAB6D457 CRC64;

  Query Match             31.4%;  Score 16;  DB 1;  Length 9;
  Best Local Similarity   33.3%;  Pred. No. 9.3e+04;
  Matches    2;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    (

Qy          4 PNAPYL 9
               || ::
Db          2 PNQDFM 7

RESULT 9
FARA_CALVO
ID   FARA_CALVO      STANDARD;      PRT;     9 AA.
AC   P41865;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   01-NOV-1995 (Rel. 32, Last annotation update)
DE   CALLIFMRFAMIDE 10.
OS   Calliphora vomitoria (Blue blowfly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha,
OC   Oestroidea; Calliphoridae; Calliphora.
OX   NCBI_TaxID=27454;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Thoracic ganglion;
RX   MEDLINE=92196111; PubMed=1549595;
RA   Duve H., Johnsen A.H., Sewell J.C., Scott A.G., Orchard I.,
RA   Rehfeld J.F., Thorpe A.;
RT   "Isolation, structure, and activity of -Phe-Met-Arg-Phe-NH2
RT   neuropeptides (designated calliFMRFamides) from the blowfly
RT   Calliphora vomitoria.";
RL   Proc. Natl. Acad. Sci. U.S.A. 89:2326-2330(1992).
CC   -!- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)
CC       FAMILY.
DR   PIR; A44787; A44787.
KW   Neuropeptide; Amidation.
FT   MOD_RES       9      9       AMIDATION.
FT   UNSURE        1      1       OR S OR A.
SQ   SEQUENCE   9 AA;  1183 MW;  29D00699CAB40457 CRC64;

  Query Match             31.4%;  Score 16;  DB 1;  Length 9,
  Best Local Similarity   33.3%;  Pred. No 9 3e+04,
  Matches    2;  Conservative    2,  Mismatches    2,  Indels    0,  Gaps
```

```
1 PNAPYL 9
  ||  ::
2 PNRDFM 7


 10
EHY
A2_TREHY      STANDARD;          PRT;      9 AA.
)0159;
1-FEB-1995 (Rel. 31, Created)
1-FEB-1995 (Rel. 31, Last sequence update)
0-MAY-2000 (Rel. 39, Last annotation update)
LAGELLAR FILAMENT OUTER LAYER PROTEIN FLAA2 (35 KDA SHEATH PROTEIN)
FRAGMENT).
LAA2.
reponema hyodysenteriae (Serpulina hyodysenteriae).
Bacteria; Spirochaetales; Brachyspiraceae; Brachyspira.
NCBI_TaxID=159;
[1]
SEQUENCE.
STRAIN=C5;
MEDLINE=93139764; PubMed=1487733;
Koopman M.B.H., Baats E., van Vorstenbosch C.J.A.H.V.,
van der Zeijst B.A.M., Kusters J.G.;
"The periplasmic flagella of Serpulina (Treponema) hyodysenteriae are
composed of two sheath proteins and three core proteins.";
  Gen. Microbiol. 138:2697-2706(1992).
!- FUNCTION: COMPONENT OF THE OUTER LAYER OF THE FLAGELLA.
-!- SUBUNIT: OUTER LAYER IS COMPOSED OF TWO SHEATH PROTEINS, FLAA1
    (44 KDA) AND FLAA2 (35 KDA) AND A CORE THAT CONTAINS THREE
    PROTEINS FLAB1 (37 KDA), FLAB2 (34 KDA) AND FLAB3 (32 KDA).
-!- SUBCELLULAR LOCATION: PERIPLASMIC FLAGELLUM.
Flagella; Periplasmic.
UNSURE         2      2
UNSURE         8      9
NON_TER        9      9
SEQUENCE   9 AA;  1129 MW;  855A19C68B4772D1 CRC64;

uery Match             31.4%;  Score 16;  DB 1;  Length 9;
est Local Similarity   66.7%;  Pred. No. 9.3e+04;
atches    2;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

   7 PYL 9
     ||:
     PYM 6

   C.                                  PRT;      9 AA.
   RN       STANDARD;
   RP     SEQU   25, Created)
   RC     TISSUE=P5, Last sequence update)
   RX     MEDLINE=930  Last annotation update)
   RA     Hughes G.J., Fr  -PAGE OF PLASMA (SPOT 11) (FRAGMENT).
   RA     Sanchez J.-C., Jame  a; Craniata; Vertebrata; Euteleostomi;
   RA     Hochstrasser D.F.;   Catarrhini; Hominidae; Homo.
   RT     "Plasma protein map: an up
   RL     Electrophoresis 13:707-714(19
   CC     -!- MISCELLANEOUS: ON THE 2D-GEL  ier F., Pasquali C.,
   CC         PROTEIN IS: 4.6, ITS MW IS: 46 K  llqvist B.,
   DR     SWISS-2DPAGE; P30089; HUMAN.
   FT     NON_TER          1       1           cing.";
                                                PI OF THIS UNKNOWN
```

```
FT   NON_TER         9      9
SQ   SEQUENCE   9 AA;  1056 MW;  26F2B1BAF769C737 CRC64;

  Query Match             31.4%;  Score 16;  DB 1;  Length 9;
  Best Local Similarity   28.6%;  Pred. No. 9.3e+04;
  Matches    2;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    (

Qy       2 MFPNAPY 8
            :||   :
Db       3 LFPXTDF 9


RESULT 12
ALL7_CARMA
ID   ALL7_CARMA      STANDARD;          PRT;      8 AA.
AC   P81809; P81810; P81804;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   CARCINUSTATIN 7 [CONTAINS: CARCINUSTATIN 6; CARCINUSTATIN 1].
OS   Carcinus maenas (Common shore crab) (Green crab).
OC   Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC   Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
OC   Eubrachyura; Portunoidea; Portunidae; Carcinus.
OX   NCBI_TaxID=6759;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Cerebral ganglion, and Thoracic ganglion;
RX   MEDLINE=98121193; PubMed=9461295;
RA   Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,
RA   Thorpe A.;
RT   "Isolation and identification of multiple neuropeptides of the
RT   allatostatin superfamily in the shore crab Carcinus maenas.";
RL   Eur. J. Biochem. 250:727-734(1997).
CC   -!- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.
CC   -!- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
KW   Neuropeptide; Amidation; Multigene family.
FT   CHAIN          1      8       CARCINUSTATIN 7.
FT   CHAIN          2      8       CARCINUSTATIN 6.
FT   CHAIN          4      8       CARCINUSTATIN 1.
FT   MOD_RES        8      8       AMIDATION.
SQ   SEQUENCE   8 AA;  825 MW;  922879CDCB4775BD CRC64;

  Query Match             29.4%;  Score 15;  DB 1;  Length 8;
  Best Local Similarity   66.7%;  Pred. No. 9.3e+04;
  Matches    2;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0,

Qy       6 APY 8
            :||
Db       2 SPY 4


RESULT 13
LMT3_LOCMI
ID   LMT3_LOCMI      STANDARD;          PRT;      9 AA.
AC   P41489;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   01-NOV-1995 (Rel. 32, Last annotation update)
DE   LOCUSTAMYOTROPIN 3 (LOM-MT-3).
OS   Locusta migratoria (Migratory locust).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Orthopteroidea; Orthoptera; Caelifera;
OC   Acridomorpha; Acridoidea; Acrididae; Locusta.
OX   NCBI_TaxID=7004;
RN   [1]
RP   SEQUENCE, AND SYNTHESIS.
RC   TISSUE=Brain;
RA   Schoofs L., Holman G.M., Hayes T.K., Nachman R.J., Kochansky J P ,
RA   de Loof A.;
RT   "Isolation, identification and synthesis of locustamyotropin III and
```

```
RT   IV, two additional neuropeptides of Locusta migratoria: members of the
RT   locustamyotropin peptide family.";
RL   Insect Biochem. Mol. Biol. 22:447-452(1992).
CC   -!- FUNCTION: POTENT MEDIATOR OF VISCERAL MUSCLE CONTRACTILE ACTIVITY
CC       (MYOTROPIC ACTIVITY).
CC   -!- SIMILARITY: BELONGS TO THE PYROKININ FAMILY.
DR   InterPro; IPR001484; -.
DR   PROSITE; PS00539; PYROKININ; 1.
KW   Neuropeptide; Amidation; Pyrokinin.
FT   MOD_RES       9      9       AMIDATION.
SQ   SEQUENCE   9 AA;  1140 MW;  D5AE1772C9D776C6 CRC64;

  Query Match             29.4%;  Score 15;  DB 1;  Length 9;
  Best Local Similarity   44.4%;  Pred. No. 9.3e+04;
  Matches    4;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy         1 RMFPNAPYL 9
               |  |  | |
Db         1 RQQPFVPRL 9


RESULT 14
SUGA_ACHDO
ID   SUGA_ACHDO     STANDARD;      PRT;     5 AA.
AC   P19991;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   01-OCT-1996 (Rel. 34, Last annotation update)
DE   SUBOESOPHAGEAL GANGLION PENTAPEPTIDE.
OS   Acheta domesticus (House cricket).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Orthopteroidea; Orthoptera; Ensifera; Gryllidae;
OC   Gryllinae; Acheta.
OX   NCBI_TaxID=6997;
RN   [1]
RP   SEQUENCE.
RA   Wicker C., Wicker C.;
RT   "Isolation and structure of a peptide isolated from the
RT   suboesophageal ganglion of Acheta domesticus (orthoptera).";
RL   Comp. Biochem. Physiol. 88C:185-187(1987).
CC   -!- SUBCELLULAR LOCATION: MAIN PEPTIDE FROM THE SUBOESOPHAGEAL
CC       GANGLIA.
DR   PIR; JS0319; JS0319.
SQ   SEQUENCE   5 AA;  476 MW;  69D76DDDDDB00000 CRC64;

  Query Match             27.5%;  Score 14;  DB 1;  Length 5;
  Best Local Similarity   66.7%;  Pred. No. 9.3e+04;
  Matches    2;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         6 APY 8
               ||:
Db         3 APF 5


RESULT 15
ALL3_CARMA
ID   ALL3_CARMA     STANDARD;      PRT;     7 AA.
AC   P81806;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   CARCINUSTATIN 3.
OS   Carcinus maenas (Common shore crab) (Green crab).
OC   Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC   Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
OC   Eubrachyura; Portunoidea; Portunidae; Carcinus.
OX   NCBI_TaxID=6759;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Cerebral ganglion, and Thoracic ganglion;
RX   MEDLINE=98121193; PubMed=9461295;
RA   Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,
RA   Thorpe A.;
RT   "Isolation and identification of multiple neuropeptides of the
RT   allatostatin superfamily in the shore crab Carcinus maenas.";
RL   Eur. J. Biochem. 250:727-734(1997).
CC   -!- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.
CC   -!- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
KW   Neuropeptide; Multigene family.
SQ   SEQUENCE   7 AA;  796 MW;  672879CDCB476B70 CRC64;

  Query Match             27.5%;  Score 14;  DB 1;  Length 7;
  Best Local Similarity  100.0%;  Pred. No. 9.3e+04;
  Matches    2;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         7 PY 8
               ||
Db         2 PY 3


Search completed: October 23, 2001, 13:34:05
Job time: 277 sec
```

```
                    GenCore version 4.5
          Copyright (c) 1993 - 2000  Compugen Ltd.

 OM protein - protein search, using sw model

 Run on:          October 23, 2001, 13:28:58 ; Search time 32.72 Seconds
                                             (without alignments)
                                             36.392 Million cell updates/sec

 Title:           US-09-625-963-1
 Perfect score:   51
 Sequence:        1 RMFPNAPYL 9

 Scoring table:   BLOSUM62
                  Gapop 10.0 , Gapext 0.5

 Searched:        425026 seqs, 132305027 residues

 Total number of hits satisfying chosen parameters:      523

 Minimum DB seq length: 0
 Maximum DB seq length: 9

 Post-processing: Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries

 Database :       SPTREMBL_16:*
               1:  sp_archea:*
               2:  sp_bacteria:*
               3:  sp_fungi:*
               4:  sp_human:*
               5:  sp_invertebrate:*
               6:  sp_mammal:*
               7:  sp_mhc:*
               8:  sp_organelle:*
               9:  sp_phage:*
              10:   sp_plant:*
              11:   sp_rodent:*
              12:   sp_unclassified:*
              13:   sp_vertebrate:*
              14:   sp_virus:*

     Pred. No. is the number of results predicted by chance to have a
     score greater than or equal to the score of the result being printed,
     and is derived by analysis of the total score distribution.

                              SUMMARIES
                   %
 Result          Query
   No.   Score   Match Length DB  ID                      Description
 ------------------------------------------------------------------------
     1     20    39.2      9   4  Q9UMF3                 Q9umf3 homo sapien
     2     17    33.3      8   4  Q9UJ50                 Q9uj50 homo sapien
     3     17    33.3      8   5  P82692                 P82692 periplaneta
     4     17    33.3      9   4  Q9UM87                 Q9um87 homo sapien
     5     17    33.3      9  10   P82429                 P82429 nicotiana t
     6     16    31.4      8   2  Q9R7T2                 Q9r7t2 escherichia
     7     16    31.4      9   3  Q9P8E5                 Q9p8e5 kluyveromyc
     8     15    29.4      8   5  Q9TWH6                 Q9twh6 perinereis
     9     15    29.4      9   4  Q9UCS8                 Q9ucs8 homo sapien
    10     15    29.4      9   4  Q9H522                 Q9h522 homo sapien
    11     15    29.4      9  10   Q9S8J8                 Q9s8j8 oryza sativ
    12     14    27.5      7   8  O98866                 O98866 spinacia ol
    13     14    27.5      7  14   Q07624                 Q07624 rous sarcom
    14     14    27.5      8   3  P82858                 P82858 puccinia re
    15     14    27.5      9   6  Q9TRU7                 Q9tru7 bos taurus
    16     14    27.5      9   8  Q9T2K9                 Q9t2k9 spinacia ol
    17     14    27.5      9  14   Q67605                 Q67605 squash leaf
    18     14    27.5      9  14   Q67606                 Q67606 squash leaf
    19     13    25.5      8   2  Q49534                 Q49534 mycoplasma
    20     13    25.5      8   5  P82618                 P82618 periplaneta
    21     13    25.5      9   2  Q99193                 Q99193 pseudomonas
    22     13    25.5      9   4  Q9UKJ6                 Q9ukj6 homo sapien
    23     13    25.5      9   4  Q9UC36                 Q9uc36 homo sapien
    24     13    25.5      9   6  Q28121                 Q28121 bos taurus
    25     13    25.5      9   6  P82926                 P82926 bos taurus
    26     13    25.5      9  14   O71066                 O71066 canine dist
    27     12    23.5      8   4  Q9P0K3                 Q9p0k3 homo sapien
    28     12    23.5      8  10   P82324                 P82324 pisum sativ
    29     12    23.5      8  11   Q9JLD7                 Q9jld7 mesocricetu
    30     12    23.5      8  11   P82598                 P82598 rattus norv
    31     12    23.5      8  14   O90345                 O90345 hepatitis g
    32     12    23.5      8  14   Q9WJ33                 Q9wj33 pseudorabie
    33     12    23.5      9   5  Q9TWV0                 Q9twv0 anthopleura
    34     12    23.5      9   8  Q9TKF2                 Q9tkf2 asteromyrtu
    35     11    21.6      7  14   Q9YQ10                 Q9yq10 porcine tra
    36     11    21.6      8   2  Q9S443                 Q9s443 pseudomonas
    37     11    21.6      8   2  Q9RQ57                 Q9rq57 buchnera ap
    38     11    21.6      8   2  Q9RQ49                 Q9rq49 buchnera ap
    39     11    21.6      8   3  P87225                 P87225 saccharomyc
    40     11    21.6      8   4  Q9UL56                 Q9ul56 homo sapien
    41     11    21.6      8   4  Q9HCQ0                 Q9hcq0 homo sapien
    42     11    21.6      8   8  Q9MSX1                 Q9msx1 jurinea hum
    43     11    21.6      8  10   Q40530                 Q40530 nicotiana t
    44     11    21.6      8  14   Q83332                 Q83332 murine hepa
    45     11    21.6      8  14   Q84156                 Q84156 orf virus (

                              ALIGNMENTS

 RESULT   1
 Q9UMF3
 ID   Q9UMF3      PRELIMINARY;      PRT;      9 AA.
 AC   Q9UMF3;
 DT   01-MAY-2000 (TrEMBLrel. 13, Created)
 DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
 DE   PD-1 PROTEIN (FRAGMENT).
 GN   PD-1.
 OS   Homo sapiens (Human).
 OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX   NCBI_TaxID=9606;
 RN   [1]
 RP   SEQUENCE FROM N.A.
 RC   TISSUE=PLACENTA;
 RX   MEDLINE=97473511; PubMed=9332365;
 RA   Finger L.R., Pu J., Wasserman R., Vibhakar R., Louie E., Hardy R.R.,
 RA   Burrows P.D., Billips L.G.;
 RT   "The human PD-1 gene: complete cDNA, genomic organization, and
 RT   developmentally regulated expression in B cell progenitors.";
 RL   Gene 197:177-187(1997).
 DR   EMBL; U64864; AAC51774.1; -.
 FT   NON_TER       9      9
 SQ   SEQUENCE   9 AA;  1067 MW;  DD4A676DC6C76046 CRC64;

   Query Match             39.2%;  Score 20;  DB 4;  Length 9;
   Best Local Similarity   60.0%;  Pred. No. 4.2e+05;
   Matches    3;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

 Qy        4 PNAPY 8
             | ||:
 Db        4 PQAPW 8

 RESULT   2
 Q9UJ50
 ID   Q9UJ50      PRELIMINARY;      PRT;      8 AA.
 AC   Q9UJ50;
 DT   01-MAY-2000 (TrEMBLrel. 13, Created)
 DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
```

```
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   LATROPHILIN-2 (FRAGMENT).
GN   LPHH1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99153747; PubMed=10030676;
RA   White G.R.M., Varley J.M., Heighway J.;
RT   "Isolation and characterisation of a human homologue of the
RT   latrophilin gene from a region of 1p31.1 implicated in breast
RT   cancer.";
RL   Oncogene 17:3513-3519(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   White G.R.M., Varley J.M., Heighway J.;
RL   Submitted (MAY-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AJ244509; CAB60204.1; -.
FT   NON_TER       1      1
FT   NON_TER       8      8
SQ   SEQUENCE   8 AA;  1025 MW;  4D59D763344AAB0B CRC64;

  Query Match             33.3%;  Score 17;  DB 4;  Length 8;
  Best Local Similarity   40.0%;  Pred. No. 4.2e+05;
  Matches    2;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        5 NAPYL 9
            | |::
Db        4 NKPFI 8


RESULT   3
P82692
ID   P82692      PRELIMINARY;      PRT;     8 AA.
AC   P82692;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   PYROKININ-2 (PEA-PK-2) (FXPRL-AMIDE).
OS   Periplaneta americana (American cockroach).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Orthopteroidea; Dictyoptera; Blattaria;
OC   Blattoidea; Blattidae; Periplaneta.
OX   NCBI_TaxID=6978;
RN   [1]
RP   SEQUENCE, FUNCTION, AND MASS SPECTROMETRY.
RC   TISSUE=CORPORA CARDIACA;
RX   MEDLINE=97353923; PubMed=9210163;
RA   Predel R., Kellner R., Kaufmann R., Penzlin H., Gaede G.;
RT   "Isolation and structural elucidation of two pyrokinins from the
RT   retrocerebral complex of the American cockroach.";
RL   Peptides 18:473-478(1997).
RN   [2]
RP   TISSUE SPECIFICITY.
RX   MEDLINE=20189894; PubMed=20189894;
RA   Predel R., Eckert M.;
RT   "Tagma-specific distribution of FXPRLamides in the nervous system of
RT   the American cockroach.";
RL   J. Comp. Neurol. 419:352-363(2000).
CC   -!- FUNCTION: MEDIATES VISCERAL MUSCLE CONTRACTILE ACTIVITY
CC       (MYOTROPIC ACTIVITY).
CC   -!- TISSUE SPECIFICITY: CORPORA CARDIACA.
CC   -!- MASS SPECTROMETRY: MW=883; METHOD=MALDI.
CC   -!- SIMILARITY: BELONGS TO THE PYROKININ FAMILY.
KW   Neuropeptide; Amidation; Pyrokinin.
FT   MOD_RES       8      8       AMIDATION.
SQ   SEQUENCE   8 AA;  884 MW;  C834176DD9D77775 CRC64;

  Query Ma[illegible]             33.3%;  Score 17;  DB 5;  Length 8;
```

```
  Best Local Similarity   66.7%;  Pred. No. 4.2e+05;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy        4 PNAPYL 9
            | || |
Db        3 PFAPRL 8


RESULT   4
Q9UM87
ID   Q9UM87      PRELIMINARY;      PRT;     9 AA.
AC   Q9UM87;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   FGFR2 PROTEIN (FRAGMENT).
GN   FGFR2.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96253074; PubMed=8676562;
RA   Wada C., Ishigaki M., Toyo-oka Y., Yamabe H., Ohnuki Y., Takada F.,
RA   Yamazaki Y., Ohtani H.;
RT   "[Nucleotide sequences at intron 6 and exon 7 junction of fibroblast
RT   growth factor receptor 2 and rapid mutational analysis in Apert
RT   syndrome].";
RL   Rinsho Byori 44:435-438(1996).
DR   EMBL; S82438; AAD14392.1; -.
FT   NON_TER       1      1
SQ   SEQUENCE   9 AA;  1103 MW;  9E4D20477401F775 CRC64;

  Query Match             33.3%;  Score 17;  DB 4;  Length 9;
  Best Local Similarity   50.0%;  Pred. No. 4.2e+05;
  Matches    3;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy        4 PNAPYL 9
            |: | |
Db        3 PHRPIL 8


RESULT   5
P82429
ID   P82429      PRELIMINARY;      PRT;     9 AA.
AC   P82429;
DT   01-JUN-2000 (TrEMBLrel. 14, Created)
DT   01-JUN-2000 (TrEMBLrel. 14, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   44 KDA CELL WALL PROTEIN (FRAGMENT).
OS   Nicotiana tabacum (Common tobacco).
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I;
OC   Solanales; Solanaceae; Nicotiana.
OX   NCBI_TaxID=4097;
RN   [1]
RP   SEQUENCE.
RC   STRAIN=CV. PETIT HAVANA;
RA   Blee K.A., Bonham V.A., Mitchell G.P., Robertson D., Slabas A.R.,
RA   Wojtaszek P., Bolwell G.P.;
RT   "Proteomic study of secondary cell wall proteins from transformed
RT   tobacco culture.";
RL   Planta 0:0-0(2000).
CC   -!- SUBCELLULAR LOCATION: CELL WALL.
CC   -!- TISSUE SPECIFICITY: XYLEM.
KW   Cell wall.
FT   NON_TER       9      9
SQ   SEQUENCE   9 AA;  986 MW;  C22CCAADC6C77776 CRC64;

  Query Match             33.3%;  Score 17;  DB 10;  Length 9;
```

```
  Best Local Similarity   50.0%;  Pred. No. 4.2e+05;
  Matches    3;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       4 PNAPYL 9
           | | :|
Db       4 PQADFL 9


RESULT   6
Q9R7T2
ID   Q9R7T2      PRELIMINARY;      PRT;     8 AA.
AC   Q9R7T2;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   HYPOTHETICAL PROTEIN HI0004 (FRAGMENT).
GN   YQFG.
OS   Escherichia coli.
OC   Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC   Escherichia.
OX   NCBI_TaxID=562;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12;
RX   MEDLINE=97061202; PubMed=8905232;
RA   Oshima T., Aiba H., Baba T., Fujita K., Hayashi K., Honjo A.,
RA   Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,
RA   Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K.,
RA   Mori H., Motomura K., Nakamura Y., Nashimoto H., Nishio Y., Saito N.,
RA   Sampei G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,
RA   Yano M., Horiuchi T.;
RT   "A 718-kb DNA Sequence of Escherichia coli K-12 Genome Corresponding
RT   to the 12.7-28.0 min Region on the Linkage Map.";
RL   DNA Res. 3:137-155(1996).
DR   EMBL; D90705; BAA35310.1; -.
FT   NON_TER       1      1
SQ   SEQUENCE   8 AA;  964 MW;  DF133B1DD04B476A CRC64;


  Query Match             31.4%;  Score 16;  DB 2;  Length 8;
  Best Local Similarity   66.7%;  Pred. No. 4.2e+05;
  Matches    2;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       7 PYL 9
           ||:
Db       2 PYI 4


RESULT   7
Q9P8E5
ID   Q9P8E5      PRELIMINARY;      PRT;     9 AA.
AC   Q9P8E5;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   HIS4 PROTEIN (FRAGMENT).
GN   HIS4.
OS   Kluyveromyces lactis (Yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Kluyveromyces.
OX   NCBI_TaxID=28985;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NRRL-Y1140;
RX   MEDLINE=99448382; PubMed=10518937;
RA   Lamas-Maceiras M., Esperanza Cerdan E., Freire-Picos M.A.;
RT   "Kluyveromyces lacxtis HIS4 transcriptional regulation: similarities
RT   and differences to Saccharomyces cerevisiae HIS4 gene.";
RL   FEBS Lett. 458:72-76(1999).
DR   EMBL; AJ238494; CAB87125.1; -.
FT   NON_TER       9      9
SQ   SEQUENCE   9 AA;  1015 MW;  5770D2D772D2D767 CRC64;
```

```
  Query Match             31.4%;  Score 16;  DB 3;  Length 9;
  Best Local Similarity   50.0%;  Pred. No. 4.2e+05;
  Matches    3;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 MFPNAP 7
           | |  |
Db       1 MLPVVP 6


RESULT   8
Q9TWH6
ID   Q9TWH6      PRELIMINARY;      PRT;     8 AA.
AC   Q9TWH6;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   BIOACTIVE PEPTIDE P4=PUTATIVE ESOPHAGEAL NEUROREGULATOR.
OS   Perinereis vancaurica.
OC   Eukaryota; Metazoa; Annelida; Polychaeta; Palpata; Aciculata;
OC   Phyllodocida; Nereididae; Perinereis.
OX   NCBI_TaxID=6355;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=95323338; PubMed=7599979;
RA   Takahashi T., Furukawa Y., Muneoka Y., Matsushima O., Ikeda T.,
RA   Fujita T., Minakata H., Nomoto K.;
RT   "Isolation and characterization of four novel bioactive peptides from
RT   a polychaete annelid, Perinereis vancaurica.";
RL   Comp. Biochem. Physiol. C,
RL   Pharmacol. Toxicol. Endocrinol. 110:297-304(1995).
SQ   SEQUENCE   8 AA;  989 MW;  954772CAA87B0B59 CRC64;


  Query Match             29.4%;  Score 15;  DB 5;  Length 8;
  Best Local Similarity   50.0%;  Pred. No. 4.2e+05;
  Matches    2;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       5 NAPY 8
           : ||
Db       5 DVPY 8


RESULT   9
Q9UCS8
ID   Q9UCS8      PRELIMINARY;      PRT;     9 AA.
AC   Q9UCS8;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   APOLIPOPROTEIN A-I.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=92075698; PubMed=1742316;
RA   Ehnholm C., Bozas S.E., Tenkanen H., Kirszbaum L., Metso J.,
RA   Murphy B., Walker I.D.;
RT   "The apolipoprotein A-I binding protein of placenta and the SP-40,40
RT   protein of human blood are different proteins which both bind to
RT   apolipoprotein A-I.";
RL   Biochim. Biophys. Acta 1086:255-260(1991).
SQ   SEQUENCE   9 AA;  981 MW;  7FE37775A6C7776B CRC64;


  Query Match             29.4%;  Score 15;  DB 4;  Length 9;
  Best Local Similarity   50.0%;  Pred. No. 4.2e+05;
  Matches    2;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       4 PNAP 7
           | :|
```

```
Db      4 PQSP 7


RESULT  10
Q9H522
ID   Q9H522      PRELIMINARY;      PRT;     9 AA.
AC   Q9H522;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   BA90M5.3 (HCG-1) (FRAGMENT).
GN   BA90M5.3.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Tromans A.;
RL   Submitted (JUL-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AL139188; CAC15103.1; -.
FT   NON_TER       9      9
SQ   SEQUENCE   9 AA;  951 MW;  96A3ADC772C455A5 CRC64;

  Query Match             29.4%;  Score 15;  DB 4;  Length 9;
  Best Local Similarity   50.0%;  Pred. No. 4.2e+05;
  Matches    3;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 MFPNAP 7
          |  | |
Db      1 MSSNVP 6


RESULT  11
Q9S8J8
ID   Q9S8J8      PRELIMINARY;      PRT;     9 AA.
AC   Q9S8J8;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   ORYZATENSIN=BIOACTIVE PEPTIDE.
OS   Oryza sativa (Rice).
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae;
OC   Oryza.
OX   NCBI_TaxID=4530;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=95102521; PubMed=7804141;
RA   Takahashi M., Moriguchi S., Yoshikawa M., Sasaki R.;
RT   "Isolation and characterization of oryzatensin: a novel bioactive
RT   peptide with ileum-contracting and immunomodulating activities derived
RT   from rice albumin.";
RL   Biochem. Mol. Biol. Int. 33:1151-1158(1994).
SQ   SEQUENCE   9 AA;  1093 MW;  0E8C67377B56877B CRC64;

  Query Match             29.4%;  Score 15;  DB 10;  Length 9;
  Best Local Similarity   66.7%;  Pred. No. 4.2e+05;
  Matches    2;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      2 MFP 4
          |:|
Db      4 MYP 6


RESULT  12
O98866
ID   O98866      PRELIMINARY;    PRT;    7 AA.
AC   O98866;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAY-1999 (TrEMBLrel. 10, Last annotation update)
DE   CYTOCHROME B/F SUBUNIT IV (FRAGMENT).
OS   Spinacia oleracea (Spinach).
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; eudicotyledons; core eudicots; Caryophyllidae;
OC   Caryophyllales; Chenopodiaceae; Spinacia.
OX   NCBI_TaxID=3562;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=86120353; PubMed=3003688;
RA   Sijben-Mueller G., Hallick R.B., Alt J., Westhoff P., Herrmann R.G.;
RT   "Spinach plastid genes coding for initiation factor IF-1, ribosomal
RT   protein S11 and RNA polymerase alpha-subunit.";
RL   Nucleic Acids Res. 14:1029-1044(1986).
DR   EMBL; X03496; CAA27215.1; -.
KW   Chloroplast.
FT   NON_TER       1      1
SQ   SEQUENCE   7 AA;  907 MW;  644729D77409C420 CRC64;

  Query Match             27.5%;  Score 14;  DB 8;  Length 7;
  Best Local Similarity   66.7%;  Pred. No. 4.2e+05;
  Matches    2;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      7 PYL 9
          |:|
Db      4 PFL 6


RESULT  13
Q07624
ID   Q07624      PRELIMINARY;      PRT;     7 AA.
AC   Q07624;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   HYPOTHETICAL PROTEIN (FRAGMENT).
OS   Rous sarcoma virus.
OC   Viruses; Retroid viruses; Retroviridae; Avian type C retroviruses.
OX   NCBI_TaxID=11886;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PRAGUE C;
RX   MEDLINE=93010967; PubMed=1327749;
RA   Donze O., Spahr P.F.;
RT   "Role of the open reading frames of Rous sarcoma virus leader RNA in
RT   translation and genome packaging.";
RL   EMBO J. 11:3747-3757(1992).
DR   EMBL; X67587; CAA47862.1; -.
KW   Hypothetical protein.
FT   NON_TER       7      7
SQ   SEQUENCE   7 AA;  672 MW;  776045A7687DD6F0 CRC64;

  Query Match             27.5%;  Score 14;  DB 14;  Length 7;
  Best Local Similarity   50.0%;  Pred. No. 4.2e+05;
  Matches    2;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      4 PNAP 7
          |: |
Db      4 PSIP 7


RESULT  14
P82858
ID   P82858      PRELIMINARY;      PRT;     8 AA.
AC   P82858;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   URICASE (EC 1.7.3.3) (URATE OXIDASE) (FRAGMENT).
OS   Puccinia recondita f. sp. triseti.
```

```
OC   Eukaryota; Fungi; Basidiomycota; Urediniomycetes; Uredinales;
OC   Pucciniaceae; Puccinia.
OX   NCBI_TaxID=142679;
RN   [1]
RP   SEQUENCE, FUNCTION, AND DEVELOPMENTAL STAGE.
RC   TISSUE=SPORE;
RA   Aguilar M., Montalbini P., Pineda M.;
RL   Submitted (NOV-2000) to the SWISS-PROT data bank.
CC   -!- FUNCTION: INVOLVED IN HOST-PARASITE RELATIONSHIP BETWEEN PLANTS
CC       AND FUNGI.
CC   -!- CATALYTIC ACTIVITY: URATE + O(2) + H(2)O = 5-HYDROXYISOURATE +
CC       H(2)O(2) (5-HYDROXYISOURATE DECOMPOSE TO FORM ALLANTOIN).
CC   -!- SUBCELLULAR LOCATION: PEROXISOMAL (BY SIMILARITY).
CC   -!- DEVELOPMENTAL STAGE: GERMINATION.
CC   -!- SIMILARITY: BELONGS TO THE URICASE FAMILY.
DR   InterPro; IPR002042; -.
DR   PROSITE; PS00366; URICASE; PARTIAL.
KW   Oxidoreductase; Purine metabolism; Peroxisome.
FT   NON_TER       8      8
SQ   SEQUENCE   8 AA;  777 MW;  98C1ADD735B9D76D CRC64;

  Query Match             27.5%;  Score 14;  DB 3;  Length 8;
  Best Local Similarity   66.7%;  Pred. No. 4.2e+05;
  Matches    2;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          6 APY 8
              ||:
Db          1 APF 3


RESULT 15
Q9TRU7
ID   Q9TRU7       PRELIMINARY;      PRT;     9 AA.
AC   Q9TRU7;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   GAP-3, GTPASE-ACTIVATING PROTEIN.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=92112868; PubMed=1309786;
RA   Nice E.C., Fabri L., Hammacher A., Holden J., Simpson R.J.,
RA   Burgess A.W.;
RT   "The purification of a Rap1 GTPase-activating protein from bovine
RT   brain cytosol.";
RL   J. Biol. Chem. 267:1546-1553(1992).
SQ   SEQUENCE   9 AA;  1063 MW;  89EDA77B47604B5A CRC64;

  Query Match             27.5%;  Score 14;  DB 6;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 4.2e+05;
  Matches    2;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          7 PY 8
              ||
Db          5 PY 6


Search completed: October 23, 2001, 13:33:46
Job time: 288 sec
```

```
                       GenCore version 4.5
               Copyright (c) 1993 - 2000  Compugen Ltd.

OM protein - protein search, using sw model

Run on:          October 23, 2001, 13:20:07 ; Search time 241.8 Seconds
                                               (without alignments)
                                               8.257 Million cell updates/sec

Title:           US-09-625-963-1
Perfect score:   51
Sequence:        1 RMFPNAPYL 9

Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched:        2803329 seqs, 221847457 residues

Total number of hits satisfying chosen parameters:      2803329

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       Pending_Patents_AA_Main:*
               1:  /cgn2_6/ptodata/2/paa/PCTUS_COMB.pep:*
               2:  /cgn2_6/ptodata/2/paa/US06_COMB.pep:*
               3:  /cgn2_6/ptodata/2/paa/US07_COMB.pep:*
               4:  /cgn2_6/ptodata/2/paa/US080_COMB.pep:*
               5:  /cgn2_6/ptodata/2/paa/US081_COMB.pep:*
               6:  /cgn2_6/ptodata/2/paa/US082_COMB.pep:*
               7:  /cgn2_6/ptodata/2/paa/US083_COMB.pep:*
               8:  /cgn2_6/ptodata/2/paa/US084_COMB.pep:*
               9:  /cgn2_6/ptodata/2/paa/US085_COMB.pep:*
               10:  /cgn2_6/ptodata/2/paa/US086_COMB.pep:*
               11:  /cgn2_6/ptodata/2/paa/US087_COMB.pep:*
               12:  /cgn2_6/ptodata/2/paa/US088_COMB.pep:*
               13:  /cgn2_6/ptodata/2/paa/US089_COMB.pep:*
               14:  /cgn2_6/ptodata/2/paa/US090_COMB.pep:*
               15:  /cgn2_6/ptodata/2/paa/US091_COMB.pep:*
               16:  /cgn2_6/ptodata/2/paa/US092_COMB.pep:*
               17:  /cgn2_6/ptodata/2/paa/US093_COMB.pep:*
               18:  /cgn2_6/ptodata/2/paa/US094_COMB.pep:*
               19:  /cgn2_6/ptodata/2/paa/US095_COMB.pep:*
               20:  /cgn2_6/ptodata/2/paa/US096_COMB.pep:*
               21:  /cgn2_6/ptodata/2/paa/US097_COMB.pep:*
               22:  /cgn2_6/ptodata/2/paa/US098_COMB.pep:*
               23:  /cgn2_6/ptodata/2/paa/US60_COMB.pep:*

     Pred. No. is the number of results predicted by chance to have a
     score greater than or equal to the score of the result being printed,
     and is derived by analysis of the total score distribution.

                                  SUMMARIES
                %
Result          Query
  No.   Score   Match Length DB  ID                     Description
-------------------------------------------------------------------------
    1      51   100.0      9  15  US-09-164-223-185         Sequence 185, App
    2      51   100.0      9  15  US-09-164-223-293         Sequence 293, App
    3      51   100.0      9  15  US-09-164-223-185         Sequence 185, App
    4      51   100.0      9  15  US-09-164-223-293         Sequence 293, App
    5      51   100.0      9  15  US-09-164-223-185         Sequence 185, App
    6      51   100.0      9  15  US-09-164-223-293         Sequence 293, App
    7      51   100.0      9  16  US-09-276-484-185         Sequence 185, App
    8      51   100.0      9  16  US-09-276-484-293         Sequence 293, App
    9      51   100.0      9  16  US-09-276-484A-185        Sequence 185, App
   10      51   100.0      9  16  US-09-276-484A-293        Sequence 293, App
```

Cheever

```
   11      51   100.0      9  19  US-09-523-419-185   Abs   Sequence 185, App
   12      51   100.0      9  19  US-09-523-419-293         Sequence 293, App
   13      51   100.0      9  20  US-09-625-963-1  Applic   Sequence 1, Appli
   14      51   100.0      9  20  US-09-679-339-185  Am     Sequence 185, App
   15      51   100.0      9  20  US-09-679-339-293         Sequence 293, App — Me
   16      51   100.0      9  20  US-09-684-361-185  RUN    Sequence 185, App
   17      51   100.0      9  20  US-09-684-361-293         Sequence 293, App
   18      51   100.0      9  20  US-09-685-830-185  RUN    Sequence 185, App
   19      51   100.0      9  20  US-09-685-830-293         Sequence 293, App
   20      51   100.0      9  21  US-09-785-019-185         Sequence 185, App
   21      51   100.0      9  21  US-09-785-019-293         Sequence 293, App
   22      51   100.0      9  21  US-09-791-477-185         Sequence 185, App
   23      51   100.0      9  21  US-09-791-477-293         Sequence 293, App
   24      51   100.0      9  22  US-09-872-832-46          Sequence 46, Appl
   25      51   100.0     23  15  US-09-164-223-2           Sequence 2, Appli
   26      51   100.0     23  15  US-09-164-223-3           Sequence 3, Appli
   27      51   100.0     23  15  US-09-164-223-2           Sequence 2, Appli
   28      51   100.0     23  15  US-09-164-223-3           Sequence 3, Appli
   29      51   100.0     23  15  US-09-164-223-2           Sequence 2, Appli
   30      51   100.0     23  15  US-09-164-223-3           Sequence 3, Appli
   31      51   100.0     23  16  US-09-276-484-2           Sequence 2, Appli
   32      51   100.0     23  16  US-09-276-484-3           Sequence 3, Appli
   33      51   100.0     23  16  US-09-276-484A-2          Sequence 2, Appli
   34      51   100.0     23  16  US-09-276-484A-3          Sequence 3, Appli
   35      51   100.0     23  19  US-09-523-419-2           Sequence 2, Appli
   36      51   100.0     23  19  US-09-523-419-3           Sequence 3, Appli
   37      51   100.0     23  20  US-09-679-339-2           Sequence 2, Appli
   38      51   100.0     23  20  US-09-679-339-3           Sequence 3, Appli
   39      51   100.0     23  20  US-09-684-361-2           Sequence 2, Appli
   40      51   100.0     23  20  US-09-684-361-3           Sequence 3, Appli
   41      51   100.0     23  20  US-09-685-830-2           Sequence 2, Appli
   42      51   100.0     23  20  US-09-685-830-3           Sequence 3, Appli
   43      51   100.0     23  21  US-09-785-019-2           Sequence 2, Appli
   44      51   100.0     23  21  US-09-785-019-3           Sequence 3, Appli
   45      51   100.0     23  21  US-09-791-477-2           Sequence 2, Appli

                                  ALIGNMENTS

RESULT   1
US-09-164-223-185
; Sequence 185, Application US/09164223
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
;  APPLICANT: Cheever, Martin A.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1 SPECIFIC IMMUNOTHERAPY
;  FILE REFERENCE: 210121.465
;  CURRENT APPLICATION NUMBER: US/09/164,223
;  CURRENT FILING DATE: 1998-09-30
;  NUMBER OF SEQ ID NOS: 320
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 185
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-164-223-185

  Query Match             100.0%;  Score 51;  DB 15;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.6e+06;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 RMFPNAPYL 9
            |||||||||
Db        1 RMFPNAPYL 9

RESULT   2
US-09-164-223-293
; Sequence 293, Application US/09164223
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
```

```
;  APPLICANT: Cheever, Martin A.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1 SPECIFIC IMMUNOTHERAPY
;  FILE REFERENCE: 210121.465
;  CURRENT APPLICATION NUMBER: US/09/164,223
;  CURRENT FILING DATE: 1998-09-30
;  NUMBER OF SEQ ID NOS: 320
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 293
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-164-223-293

  Query Match             100.0%;  Score 51;  DB 15;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.6e+06;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 RMFPNAPYL 9
            |||||||||
Db        1 RMFPNAPYL 9

RESULT   3
US-09-164-223-185
; Sequence 185, Application US/09164223A
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
;  APPLICANT: Cheever, Martin A.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1 SPECIFIC IMMUNOTHERAPY
;  FILE REFERENCE: 210121.465
;  CURRENT APPLICATION NUMBER: US/09/164,223A
;  CURRENT FILING DATE: 1998-09-30
;  NUMBER OF SEQ ID NOS: 320
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 185
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-164-223-185

  Query Match             100.0%;  Score 51;  DB 15;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.6e+06;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 RMFPNAPYL 9
            |||||||||
Db        1 RMFPNAPYL 9

RESULT   4
US-09-164-223-293
; Sequence 293, Application US/09164223A
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
;  APPLICANT: Cheever, Martin A.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1 SPECIFIC IMMUNOTHERAPY
;  FILE REFERENCE: 210121.465
;  CURRENT APPLICATION NUMBER: US/09/164,223A
;  CURRENT FILING DATE: 1998-09-30
;  NUMBER OF SEQ ID NOS: 320
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 293
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-164-223-293

  Query Match             100.0%;  Score 51;  DB 15;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.6e+06;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 RMFPNAPYL 9
            |||||||||
Db        1 RMFPNAPYL 9

RESULT   5
US-09-164-223-185
; Sequence 185, Application US/09164223B
; GENERAL INFORMATION:
;  APPLICANT:  Gaiger, Alexander
;  APPLICANT:  Cheever, Martin A.
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR WT1 SPECIFIC IMMUNOTHERAPY
;  FILE REFERENCE:  210121.465
;  CURRENT APPLICATION NUMBER: US/09/164,223B
;  CURRENT FILING DATE:  1998-09-30
;  NUMBER OF SEQ ID NOS:  320
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 185
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-164-223-185

  Query Match             100.0%;  Score 51;  DB 15;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.6e+06;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 RMFPNAPYL 9
            |||||||||
Db        1 RMFPNAPYL 9

RESULT   6
US-09-164-223-293
; Sequence 293, Application US/09164223B
; GENERAL INFORMATION:
;  APPLICANT:  Gaiger, Alexander
;  APPLICANT:  Cheever, Martin A.
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR WT1 SPECIFIC IMMUNOTHERAPY
;  FILE REFERENCE:  210121.465
;  CURRENT APPLICATION NUMBER: US/09/164,223B
;  CURRENT FILING DATE:  1998-09-30
;  NUMBER OF SEQ ID NOS:  320
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 293
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-164-223-293

  Query Match             100.0%;  Score 51;  DB 15;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.6e+06;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 RMFPNAPYL 9
            |||||||||
Db        1 RMFPNAPYL 9

RESULT   7
US-09-276-484-185
; Sequence 185, Application US/09276484
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
;  APPLICANT: Cheever, Martin A.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1
;  TITLE OF INVENTION: SPECIFIC IMMUNOTHERAPY
;  FILE REFERENCE: 210121.465C1
;  CURRENT APPLICATION NUMBER: US/09/276,484
;  CURRENT FILING DATE: 1999-03-25
```

```
;  NUMBER OF SEQ ID NOS: 326
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 185
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Homo sapien
US-09-276-484-185


  Query Match            100.0%;  Score 51;  DB 16;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.6e+06;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 RMFPNAPYL 9
            |||||||||
Db        1 RMFPNAPYL 9

RESULT   8
US-09-276-484-293
; Sequence 293, Application US/09276484
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
;  APPLICANT: Cheever, Martin A.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1
;  TITLE OF INVENTION: SPECIFIC IMMUNOTHERAPY
;  FILE REFERENCE: 210121.465C1
;  CURRENT APPLICATION NUMBER: US/09/276,484
;  CURRENT FILING DATE: 1999-03-25
;  NUMBER OF SEQ ID NOS: 326
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 293
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-276-484-293


  Query Match            100.0%;  Score 51;  DB 16;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.6e+06;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 RMFPNAPYL 9
            |||||||||
Db        1 RMFPNAPYL 9

RESULT   9
US-09-276-484A-185
; Sequence 185, Application US/09276484A
; GENERAL INFORMATION:
;  APPLICANT:  Gaiger, Alexander
;  APPLICANT:  Cheever, Martin A.
;  TITLE OF INVENTION:  COMPOSITONS AND METHODS FOR WT1 SPECIFIC
;  TITLE OF INVENTION:  IMMUNOTHERAPY
;  FILE REFERENCE:  210121.465C1
;  CURRENT APPLICATION NUMBER: US/09/276,484A
;  CURRENT FILING DATE:  1999-03-25
;  PRIOR APPLICATION NUMBER:  US 09/164,223
;  PRIOR FILING DATE:  1999-09-30
;  NUMBER OF SEQ ID NOS:  326
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 185
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Homo sapien
US-09-276-484A-185


  Query Match            100.0%;  Score 51;  DB 16;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.6e+06;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy        1 RMFPNAPYL 9
            |||||||||
Db        1 RMFPNAPYL 9

RESULT  10
US-09-276-484A-293
; Sequence 293, Application US/09276484A
; GENERAL INFORMATION:
;  APPLICANT:  Gaiger, Alexander
;  APPLICANT:  Cheever, Martin A.
;  TITLE OF INVENTION:  COMPOSITONS AND METHODS FOR WT1 SPECIFIC
;  TITLE OF INVENTION:  IMMUNOTHERAPY
;  FILE REFERENCE:  210121.465C1
;  CURRENT APPLICATION NUMBER: US/09/276,484A
;  CURRENT FILING DATE:  1999-03-25
;  PRIOR APPLICATION NUMBER:  US 09/164,223
;  PRIOR FILING DATE:  1999-09-30
;  NUMBER OF SEQ ID NOS:  326
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 293
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-276-484A-293


  Query Match            100.0%;  Score 51;  DB 16;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.6e+06;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 RMFPNAPYL 9
            |||||||||
Db        1 RMFPNAPYL 9

RESULT  11
US-09-523-419-185
; Sequence 185, Application US/09523419
; GENERAL INFORMATION:
;  APPLICANT: Hendrickson, Ronald C.
;  APPLICANT: Probst, Peter
;  APPLICANT: Reed, Steven G.
;  TITLE OF INVENTION: METHODS FOR THE DIAGNOSIS AND THERAPY
;  TITLE OF INVENTION: OF CHLAMYDIA INFECTION
;  FILE REFERENCE: 210121.492
;  CURRENT APPLICATION NUMBER: US/09/523,419
;  CURRENT FILING DATE: 2000-03-10
;  NUMBER OF SEQ ID NOS: 326
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 185
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Homo sapien
US-09-523-419-185


  Query Match            100.0%;  Score 51;  DB 19;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.6e+06;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 RMFPNAPYL 9
            |||||||||
Db        1 RMFPNAPYL 9

RESULT  12
US-09-523-419-293
; Sequence 293, Application US/09523419
; GENERAL INFORMATION:
;  APPLICANT: Hendrickson, Ronald C.
;  APPLICANT: Probst, Peter
;  APPLICANT: Reed, Steven G.
```

```
;  TITLE OF INVENTION: METHODS FOR THE DIAGNOSIS AND THERAPY
;  TITLE OF INVENTION: OF CHLAMYDIA INFECTION
;  FILE REFERENCE: 210121.492
;  CURRENT APPLICATION NUMBER: US/09/523,419
;  CURRENT FILING DATE: 2000-03-10
;  NUMBER OF SEQ ID NOS: 326
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 293
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-523-419-293

  Query Match             100.0%;  Score 51;  DB 19;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.6e+06;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RMFPNAPYL 9
           |||||||||
Db       1 RMFPNAPYL 9

RESULT  13
US-09-625-963-1
; Sequence 1, Application US/09625963
; GENERAL INFORMATION:
;  APPLICANT:  Strauss, Hans Josef
;  APPLICANT:  Gao, Liquan
;  TITLE OF INVENTION:  Immunotherapeutic Methods Using Epitopes of WT-1 and
;  TITLE OF INVENTION:  GATA-1
;  FILE REFERENCE:  ICI 101
;  CURRENT APPLICATION NUMBER: US/09/625,963
;  CURRENT FILING DATE:  2000-07-26
;  PRIOR APPLICATION NUMBER:  PCT/GB99/03572
;  PRIOR FILING DATE:  1999-11-02
;  PRIOR APPLICATION NUMBER:  GB9823897.5
;  PRIOR FILING DATE:  1998-11-02
;  NUMBER OF SEQ ID NOS:  5
;  SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-625-963-1

  Query Match             100.0%;  Score 51;  DB 20;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.6e+06;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RMFPNAPYL 9
           |||||||||
Db       1 RMFPNAPYL 9

RESULT  14
US-09-679-339-185
; Sequence 185, Application US/09679339
; GENERAL INFORMATION:
;  APPLICANT:  Skeiky, Yasir
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Cheever, Martin
;  APPLICANT:  Reed, Steven G.
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR WT1
;  TITLE OF INVENTION:  SPECIFIC IMMUNOTHERAPY
;  FILE REFERENCE:  210121.533
;  CURRENT APPLICATION NUMBER: US/09/679,339
;  CURRENT FILING DATE:  2000-10-04
;  NUMBER OF SEQ ID NOS:  357
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 185
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Homo sapien
US-09-679-339-185

  Query Match             100.0%;  Score 51;  DB 20;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.6e+06;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RMFPNAPYL 9
           |||||||||
Db       1 RMFPNAPYL 9

RESULT  15
US-09-679-339-293
; Sequence 293, Application US/09679339
; GENERAL INFORMATION:
;  APPLICANT:  Skeiky, Yasir
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Cheever, Martin
;  APPLICANT:  Reed, Steven G.
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR WT1
;  TITLE OF INVENTION:  SPECIFIC IMMUNOTHERAPY
;  FILE REFERENCE:  210121.533
;  CURRENT APPLICATION NUMBER: US/09/679,339
;  CURRENT FILING DATE:  2000-10-04
;  NUMBER OF SEQ ID NOS:  357
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 293
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-679-339-293

  Query Match             100.0%;  Score 51;  DB 20;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.6e+06;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RMFPNAPYL 9
           |||||||||
Db       1 RMFPNAPYL 9


Search completed: October 23, 2001, 13:26:59
Job time: 412 sec
```

```
                    GenCore version 4.5
              Copyright (c) 1993 - 2000  Compugen Ltd.

OM protein - protein search, using sw model

Run on:          October 23, 2001, 13:20:07 ; Search time 39.01 Seconds
                                              (without alignments)
                                              11.888 Million cell updates/sec

Title:           US-09-625-963-1
Perfect score:   51
Sequence:        1 RMFPNAPYL 9

Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched:        339663 seqs, 51527020 residues

Total number of hits satisfying chosen parameters:      339663

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       Pending_Patents_AA_New:*
                1:  /cgn2_6/ptodata/2/paa/PCT_NEW_COMB.pep:*
                2:  /cgn2_6/ptodata/2/paa/US06_NEW_COMB.pep:*
                3:  /cgn2_6/ptodata/2/paa/US07_NEW_COMB.pep:*
                4:  /cgn2_6/ptodata/2/paa/US08_NEW_COMB.pep:*
                5:  /cgn2_6/ptodata/2/paa/US09_NEW_COMB.pep:*
                6:  /cgn2_6/ptodata/2/paa/US60_NEW_COMB.pep:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                              SUMMARIES
                 %
Result         Query
  No.  Score   Match Length DB  ID                       Description
-------------------------------------------------------------------------
    1     51   100.0      9  5  US-09-938-864-185          Sequence 185, App
    2     51   100.0      9  5  US-09-938-864-293          Sequence 293, App
    3     51   100.0     23  5  US-09-938-864-2            Sequence 2, Appli
    4     51   100.0     23  5  US-09-938-864-3            Sequence 3, Appli
    5     51   100.0    152  5  US-09-938-864-343          Sequence 343, App
    6     51   100.0    154  5  US-09-929-315-5            Sequence 5, Appli
    7     51   100.0    214  5  US-09-938-864-395          Sequence 395, App
    8     51   100.0    256  5  US-09-938-864-335          Sequence 335, App
    9     51   100.0    344  5  US-09-938-864-391          Sequence 391, App
   10     51   100.0    345  5  US-09-929-315-2            Sequence 2, Appli
   11     51   100.0    362  5  US-09-938-864-394          Sequence 394, App
   12     51   100.0    369  5  US-09-938-864-346          Sequence 346, App
   13     51   100.0    410  5  US-09-938-864-333          Sequence 333, App
   14     51   100.0    420  5  US-09-938-864-393          Sequence 393, App
   15     51   100.0    428  5  US-09-938-864-405          Sequence 405, App
   16     51   100.0    429  5  US-09-938-864-408          Sequence 408, App
   17     51   100.0    449  5  US-09-538-092-960          Sequence 960, App
   18     51   100.0    449  5  US-09-929-315-4            Sequence 4, Appli
   19     51   100.0    449  5  US-09-929-315-6            Sequence 6, Appli
   20     51   100.0    449  5  US-09-938-864-319          Sequence 319, App
   21     51   100.0    449  5  US-09-938-864-320          Sequence 320, App
   22     51   100.0    449  5  US-09-938-864-404          Sequence 404, App
   23     51   100.0    495  5  US-09-938-864-409          Sequence 409, App
   24     51   100.0    504  5  US-09-938-864-410          Sequence 410, App
   25     51   100.0    568  5  US-09-938-864-392          Sequence 392, App
   26     51   100.0    591  5  US-09-938-864-334          Sequence 334, App
   27     47    92.2      9  5  US-09-938-864-38           Sequence 38, Appl
```

```
   28     46    90.2      9  5  US-09-938-864-324          Sequence 324, App
   29     41    80.4      9  5  US-09-938-864-79           Sequence 79, Appl
   30     41    80.4      9  5  US-09-938-864-267          Sequence 267, App
   31     40    78.4      9  5  US-09-938-864-323          Sequence 323, App
   32     40    78.4    417  5  US-09-938-864-407          Sequence 407, App
   33     37    72.5    398  5  US-09-570-581A-1470        Sequence 1470, Ap
   34     36    70.6     76  1  PCT-US01-08631-54141       Sequence 54141, A
   35     36    70.6     82  1  PCT-US01-08631-54140       Sequence 54140, A
   36     36    70.6    118  5  US-09-834-366-26444        Sequence 26444, A
   37     36    70.6    118  5  US-09-890-688-134          Sequence 134, App
   38     36    70.6    134  5  US-09-834-366-24922        Sequence 24922, A
   39     36    70.6    159  1  PCT-US01-08631-54144       Sequence 54144, A
   40     36    70.6    175  5  US-09-760-475-2246         Sequence 2246, Ap
   41     36    70.6    183  5  US-09-760-475-2254         Sequence 2254, Ap
   42     36    70.6    210  1  PCT-US01-08631-54147       Sequence 54147, A
   43     36    70.6    210  5  US-09-760-475-2252         Sequence 2252, Ap
   44     36    70.6    228  5  US-09-760-475-3335         Sequence 3335, Ap
   45     36    70.6    229  5  US-09-760-475-2251         Sequence 2251, Ap

                              ALIGNMENTS


RESULT   1
US-09-938-864-185
; Sequence 185, Application US/09938864
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Smithgall, Molly
;  APPLICANT:  Moulton, Gus
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Sleath, Paul R.
;  APPLICANT:  Mossman, Sally
;  APPLICANT:  Evans, Lawrence
;  APPLICANT:  Spies, A. Gregory
;  APPLICANT:  Boydston, Jeremy
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1
;  TITLE OF INVENTION:  SPECIFIC IMMUNOTHERAPY
;  FILE REFERENCE: 210121.465C5
;  CURRENT APPLICATION NUMBER: US/09/938,864
;  CURRENT FILING DATE:  2001-08-24
;  NUMBER OF SEQ ID NOS: 413
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 185
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Homo sapien
US-09-938-864-185


  Query Match             100.0%;  Score 51;  DB 5;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.9e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RMFPNAPYL 9
           |||||||||
Db       1 RMFPNAPYL 9


RESULT   2
US-09-938-864-293
; Sequence 293, Application US/09938864
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Smithgall, Molly
;  APPLICANT:  Moulton, Gus
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Sleath, Paul R.
;  APPLICANT:  Mossman, Sally
;  APPLICANT:  Evans, Lawrence
;  APPLICANT:  Spies, A. Gregory
```

```
;  APPLICANT:  Boydston, Jeremy
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1
;  TITLE OF INVENTION:  SPECIFIC IMMUNOTHERAPY
;  FILE REFERENCE: 210121.465C5
;  CURRENT APPLICATION NUMBER: US/09/938,864
;  CURRENT FILING DATE:  2001-08-24
;  NUMBER OF SEQ ID NOS: 413
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 293
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-938-864-293

  Query Match             100.0%;  Score 51;  DB 5;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.9e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RMFPNAPYL 9
           |||||||||
Db       1 RMFPNAPYL 9

RESULT   3
US-09-938-864-2
; Sequence 2, Application US/09938864
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Smithgall, Molly
;  APPLICANT:  Moulton, Gus
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Sleath, Paul R.
;  APPLICANT:  Mossman, Sally
;  APPLICANT:  Evans, Lawrence
;  APPLICANT:  Spies, A. Gregory
;  APPLICANT:  Boydston, Jeremy
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1
;  TITLE OF INVENTION:  SPECIFIC IMMUNOTHERAPY
;  FILE REFERENCE: 210121.465C5
;  CURRENT APPLICATION NUMBER: US/09/938,864
;  CURRENT FILING DATE:  2001-08-24
;  NUMBER OF SEQ ID NOS: 413
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 2
;   LENGTH: 23
;   TYPE: PRT
;   ORGANISM: Homo sapien
US-09-938-864-2

  Query Match             100.0%;  Score 51;  DB 5;  Length 23;
  Best Local Similarity   100.0%;  Pred. No. 0.011;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RMFPNAPYL 9
           |||||||||
Db      10 RMFPNAPYL 18

RESULT   4
US-09-938-864-3
; Sequence 3, Application US/09938864
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Smithgall, Molly
;  APPLICANT:  Moulton, Gus
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Sleath, Paul R.
;  APPLICANT:  Mossman, Sally
;  APPLICANT:  Evans, Lawrence
```

```
;  APPLICANT:  Spies, A. Gregory
;  APPLICANT:  Boydston, Jeremy
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1
;  TITLE OF INVENTION:  SPECIFIC IMMUNOTHERAPY
;  FILE REFERENCE: 210121.465C5
;  CURRENT APPLICATION NUMBER: US/09/938,864
;  CURRENT FILING DATE:  2001-08-24
;  NUMBER OF SEQ ID NOS: 413
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 3
;   LENGTH: 23
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-938-864-3

  Query Match             100.0%;  Score 51;  DB 5;  Length 23;
  Best Local Similarity   100.0%;  Pred. No. 0.011;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RMFPNAPYL 9
           |||||||||
Db      10 RMFPNAPYL 18

RESULT   5
US-09-938-864-343
; Sequence 343, Application US/09938864
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Smithgall, Molly
;  APPLICANT:  Moulton, Gus
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Sleath, Paul R.
;  APPLICANT:  Mossman, Sally
;  APPLICANT:  Evans, Lawrence
;  APPLICANT:  Spies, A. Gregory
;  APPLICANT:  Boydston, Jeremy
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1
;  TITLE OF INVENTION:  SPECIFIC IMMUNOTHERAPY
;  FILE REFERENCE: 210121.465C5
;  CURRENT APPLICATION NUMBER: US/09/938,864
;  CURRENT FILING DATE:  2001-08-24
;  NUMBER OF SEQ ID NOS: 413
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 343
;   LENGTH: 152
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-938-864-343

  Query Match             100.0%;  Score 51;  DB 5;  Length 152;
  Best Local Similarity   100.0%;  Pred. No. 0.069;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RMFPNAPYL 9
           |||||||||
Db      51 RMFPNAPYL 59

RESULT   6
US-09-929-315-5
; Sequence 5, Application US/09929315
; GENERAL INFORMATION:
;  APPLICANT: Call, Katherine M.
;  APPLICANT:  Glaser, Thomas M.
;  APPLICANT:  Ito, Caryn Y.
;  APPLICANT:  Buckler, Alan J.
;  APPLICANT:  Pelletier, Jerry
;  APPLICANT:  Haber, Daniel A.
;  APPLICANT:  Rose, Elise A.
```

```
;  APPLICANT:  Housman, David E.
;  APPLICANT:  Bruening, Wendy
;  APPLICANT:  Darveau, Andre
;  TITLE OF INVENTION: Localization and Characterization of the
;  TITLE OF INVENTION:  Wilms' Tumor Gene
;  FILE REFERENCE: 0050.1312-013
;  CURRENT APPLICATION NUMBER: US/09/929,315
;  CURRENT FILING DATE:  2001-08-14
;  PRIOR APPLICATION NUMBER: US 09/037,179
;  PRIOR FILING DATE: 1998-03-09
;  PRIOR APPLICATION NUMBER: US 08/102,942
;  PRIOR FILING DATE: 1993-08-02
;  PRIOR APPLICATION NUMBER: US 07/614,161
;  PRIOR FILING DATE: 1990-11-13
;  PRIOR APPLICATION NUMBER: US 07/435,780
;  PRIOR FILING DATE: 1989-11-13
;  PRIOR APPLICATION NUMBER: US 07/795,323
;  PRIOR FILING DATE: 1991-11-20
;  NUMBER OF SEQ ID NOS: 21
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5
;   LENGTH: 154
;   TYPE: PRT
;   ORGANISM: Homo sapien
US-09-929-315-5

  Query Match             100.0%;  Score 51;  DB 5;  Length 154;
  Best Local Similarity   100.0%;  Pred. No. 0.069;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RMFPNAPYL 9
           |||||||||
Db      40 RMFPNAPYL 48

RESULT   7
US-09-938-864-395
; Sequence 395, Application US/09938864
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Smithgall, Molly
;  APPLICANT:  Moulton, Gus
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Sleath, Paul R.
;  APPLICANT:  Mossman, Sally
;  APPLICANT:  Evans, Lawrence
;  APPLICANT:  Spies, A. Gregory
;  APPLICANT:  Boydston, Jeremy
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1
;  TITLE OF INVENTION:  SPECIFIC IMMUNOTHERAPY
;  FILE REFERENCE: 210121.465C5
;  CURRENT APPLICATION NUMBER: US/09/938,864
;  CURRENT FILING DATE:  2001-08-24
;  NUMBER OF SEQ ID NOS: 413
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 395
;   LENGTH: 214
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-938-864-395

  Query Match             100.0%;  Score 51;  DB 5;  Length 214;
  Best Local Similarity   100.0%;  Pred. No. 0.096;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RMFPNAPYL 9
           |||||||||
Db      59 RMFPNAPYL 67
```

```
RESULT   8
US-09-938-864-335
; Sequence 335, Application US/09938864
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Smithgall, Molly
;  APPLICANT:  Moulton, Gus
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Sleath, Paul R.
;  APPLICANT:  Mossman, Sally
;  APPLICANT:  Evans, Lawrence
;  APPLICANT:  Spies, A. Gregory
;  APPLICANT:  Boydston, Jeremy
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1
;  TITLE OF INVENTION:  SPECIFIC IMMUNOTHERAPY
;  FILE REFERENCE: 210121.465C5
;  CURRENT APPLICATION NUMBER: US/09/938,864
;  CURRENT FILING DATE:  2001-08-24
;  NUMBER OF SEQ ID NOS: 413
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 335
;   LENGTH: 256
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-938-864-335

  Query Match             100.0%;  Score 51;  DB 5;  Length 256;
  Best Local Similarity   100.0%;  Pred. No. 0.11;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RMFPNAPYL 9
           |||||||||
Db     133 RMFPNAPYL 141

RESULT   9
US-09-938-864-391
; Sequence 391, Application US/09938864
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Smithgall, Molly
;  APPLICANT:  Moulton, Gus
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Sleath, Paul R.
;  APPLICANT:  Mossman, Sally
;  APPLICANT:  Evans, Lawrence
;  APPLICANT:  Spies, A. Gregory
;  APPLICANT:  Boydston, Jeremy
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1
;  TITLE OF INVENTION:  SPECIFIC IMMUNOTHERAPY
;  FILE REFERENCE: 210121.465C5
;  CURRENT APPLICATION NUMBER: US/09/938,864
;  CURRENT FILING DATE:  2001-08-24
;  NUMBER OF SEQ ID NOS: 413
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 391
;   LENGTH: 344
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-938-864-391

  Query Match             100.0%;  Score 51;  DB 5;  Length 344;
  Best Local Similarity   100.0%;  Pred. No. 0.15;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RMFPNAPYL 9
           |||||||||
Db     189 RMFPNAPYL 197
```

```
RESULT  10
US-09-929-315-2
; Sequence 2, Application US/09929315
; GENERAL INFORMATION:
;  APPLICANT: Call, Katherine M.
;  APPLICANT:  Glaser, Thomas M.
;  APPLICANT:  Ito, Caryn Y.
;  APPLICANT:  Buckler, Alan J.
;  APPLICANT:  Pelletier, Jerry
;  APPLICANT:  Haber, Daniel A.
;  APPLICANT:  Rose, Elise A.
;  APPLICANT:  Housman, David E.
;  APPLICANT:  Bruening, Wendy
;  APPLICANT:  Darveau, Andre
;  TITLE OF INVENTION: Localization and Characterization of the
;  TITLE OF INVENTION:  Wilms' Tumor Gene
;  FILE REFERENCE: 0050.1312-013
;  CURRENT APPLICATION NUMBER: US/09/929,315
;  CURRENT FILING DATE:  2001-08-14
;  PRIOR APPLICATION NUMBER: US 09/037,179
;  PRIOR FILING DATE: 1998-03-09
;  PRIOR APPLICATION NUMBER: US 08/102,942
;  PRIOR FILING DATE: 1993-08-02
;  PRIOR APPLICATION NUMBER: US 07/614,161
;  PRIOR FILING DATE: 1990-11-13
;  PRIOR APPLICATION NUMBER: US 07/435,780
;  PRIOR FILING DATE: 1989-11-13
;  PRIOR APPLICATION NUMBER: US 07/795,323
;  PRIOR FILING DATE: 1991-11-20
;  NUMBER OF SEQ ID NOS: 21
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
;   LENGTH: 345
;   TYPE: PRT
;   ORGANISM: Homo sapien
US-09-929-315-2


  Query Match            100.0%;  Score 51;  DB 5;  Length 345;
  Best Local Similarity   100.0%;  Pred. No. 0.15;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 RMFPNAPYL 9
          |||||||||
Db     42 RMFPNAPYL 50

RESULT  11
US-09-938-864-394
; Sequence 394, Application US/09938864
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Smithgall, Molly
;  APPLICANT:  Moulton, Gus
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Sleath, Paul R.
;  APPLICANT:  Mossman, Sally
;  APPLICANT:  Evans, Lawrence
;  APPLICANT:  Spies, A. Gregory
;  APPLICANT:  Boydston, Jeremy
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1
;  TITLE OF INVENTION:  SPECIFIC IMMUNOTHERAPY
;  FILE REFERENCE: 210121.465C5
;  CURRENT APPLICATION NUMBER: US/09/938,864
;  CURRENT FILING DATE:  2001-08-24
;  NUMBER OF SEQ ID NOS: 413
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 394
;   LENGTH: 362
;   TYPE: PRT
;   ORGANISM: Homo sapiens
```

```
US-09-938-864-394

  Query Match            100.0%;  Score 51;  DB 5;  Length 362;
  Best Local Similarity   100.0%;  Pred. No. 0.16;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 RMFPNAPYL 9
          |||||||||
Db     59 RMFPNAPYL 67

RESULT  12
US-09-938-864-346
; Sequence 346, Application US/09938864
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Smithgall, Molly
;  APPLICANT:  Moulton, Gus
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Sleath, Paul R.
;  APPLICANT:  Mossman, Sally
;  APPLICANT:  Evans, Lawrence
;  APPLICANT:  Spies, A. Gregory
;  APPLICANT:  Boydston, Jeremy
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1
;  TITLE OF INVENTION:  SPECIFIC IMMUNOTHERAPY
;  FILE REFERENCE: 210121.465C5
;  CURRENT APPLICATION NUMBER: US/09/938,864
;  CURRENT FILING DATE:  2001-08-24
;  NUMBER OF SEQ ID NOS: 413
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 346
;   LENGTH: 369
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-938-864-346


  Query Match            100.0%;  Score 51;  DB 5;  Length 369;
  Best Local Similarity   100.0%;  Pred. No. 0.16;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 RMFPNAPYL 9
          |||||||||
Db     66 RMFPNAPYL 74

RESULT  13
US-09-938-864-333
; Sequence 333, Application US/09938864
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Smithgall, Molly
;  APPLICANT:  Moulton, Gus
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Sleath, Paul R.
;  APPLICANT:  Mossman, Sally
;  APPLICANT:  Evans, Lawrence
;  APPLICANT:  Spies, A. Gregory
;  APPLICANT:  Boydston, Jeremy
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1
;  TITLE OF INVENTION:  SPECIFIC IMMUNOTHERAPY
;  FILE REFERENCE: 210121.465C5
;  CURRENT APPLICATION NUMBER: US/09/938,864
;  CURRENT FILING DATE:  2001-08-24
;  NUMBER OF SEQ ID NOS: 413
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 333
;   LENGTH: 410
;   TYPE: PRT
```

```
;   ORGANISM: Homo sapiens
US-09-938-864-333


  Query Match             100.0%;  Score 51;  DB 5;  Length 410;
  Best Local Similarity   100.0%;  Pred. No. 0.18;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 RMFPNAPYL 9
          |||||||||
Db    287 RMFPNAPYL 295


RESULT  14
US-09-938-864-393
; Sequence 393, Application US/09938864
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Smithgall, Molly
;  APPLICANT:  Moulton, Gus
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Sleath, Paul R.
;  APPLICANT:  Mossman, Sally
;  APPLICANT:  Evans, Lawrence
;  APPLICANT:  Spies, A. Gregory
;  APPLICANT:  Boydston, Jeremy
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1
;  TITLE OF INVENTION:  SPECIFIC IMMUNOTHERAPY
;  FILE REFERENCE: 210121.465C5
;  CURRENT APPLICATION NUMBER: US/09/938,864
;  CURRENT FILING DATE:  2001-08-24
;  NUMBER OF SEQ ID NOS: 413
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 393
;   LENGTH: 420
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-938-864-393


  Query Match             100.0%;  Score 51;  DB 5;  Length 420;
  Best Local Similarity   100.0%;  Pred. No. 0.19;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 RMFPNAPYL 9
          |||||||||
Db    265 RMFPNAPYL 273


RESULT  15
US-09-938-864-405
; Sequence 405, Application US/09938864
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Smithgall, Molly
;  APPLICANT:  Moulton, Gus
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Sleath, Paul R.
;  APPLICANT:  Mossman, Sally
;  APPLICANT:  Evans, Lawrence
;  APPLICANT:  Spies, A. Gregory
;  APPLICANT:  Boydston, Jeremy
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1
;  TITLE OF INVENTION:  SPECIFIC IMMUNOTHERAPY
;  FILE REFERENCE: 210121.465C5
;  CURRENT APPLICATION NUMBER: US/09/938,864
;  CURRENT FILING DATE:  2001-08-24
;  NUMBER OF SEQ ID NOS: 413
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 405
;   LENGTH: 428
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-938-864-405


  Query Match             100.0%;  Score 51;  DB 5;  Length 428;
  Best Local Similarity   100.0%;  Pred. No. 0.19;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 RMFPNAPYL 9
          |||||||||
Db    126 RMFPNAPYL 134


Search completed: October 23, 2001, 13:22:51
Job time: 164 sec
```

```
                    GenCore version 4.5
             Copyright (c) 1993 - 2000  Compugen Ltd.

OM protein - protein search, using sw model

Run on:         October 23, 2001, 13:20:07 ; Search time 46.09 Seconds
                                             (without alignments)
                                             11.838 Million cell updates/sec

Title:          US-09-625-963-1
Perfect score:  51
Sequence:       1 RMFPNAPYL 9

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched:       412676 seqs, 60623988 residues

Total number of hits satisfying chosen parameters:      412676

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       A_Geneseq_0601:*
               1:  /SIDS8/gcgdata/geneseq/geneseqp/AA1980.DAT:*
               2:  /SIDS8/gcgdata/geneseq/geneseqp/AA1981.DAT:*
               3:  /SIDS8/gcgdata/geneseq/geneseqp/AA1982.DAT:*
               4:  /SIDS8/gcgdata/geneseq/geneseqp/AA1983.DAT:*
               5:  /SIDS8/gcgdata/geneseq/geneseqp/AA1984.DAT:*
               6:  /SIDS8/gcgdata/geneseq/geneseqp/AA1985.DAT:*
               7:  /SIDS8/gcgdata/geneseq/geneseqp/AA1986.DAT:*
               8:  /SIDS8/gcgdata/geneseq/geneseqp/AA1987.DAT:*
               9:  /SIDS8/gcgdata/geneseq/geneseqp/AA1988.DAT:*
              10:  /SIDS8/gcgdata/geneseq/geneseqp/AA1989.DAT:*
              11:  /SIDS8/gcgdata/geneseq/geneseqp/AA1990.DAT:*
              12:  /SIDS8/gcgdata/geneseq/geneseqp/AA1991.DAT:*
              13:  /SIDS8/gcgdata/geneseq/geneseqp/AA1992.DAT:*
              14:  /SIDS8/gcgdata/geneseq/geneseqp/AA1993.DAT:*
              15:  /SIDS8/gcgdata/geneseq/geneseqp/AA1994.DAT:*
              16:  /SIDS8/gcgdata/geneseq/geneseqp/AA1995.DAT:*
              17:  /SIDS8/gcgdata/geneseq/geneseqp/AA1996.DAT:*
              18:  /SIDS8/gcgdata/geneseq/geneseqp/AA1997.DAT:*
              19:  /SIDS8/gcgdata/geneseq/geneseqp/AA1998.DAT:*
              20:  /SIDS8/gcgdata/geneseq/geneseqp/AA1999.DAT:*
              21:  /SIDS8/gcgdata/geneseq/geneseqp/AA2000.DAT:*
              22:  /SIDS8/gcgdata/geneseq/geneseqp/AA2001.DAT:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                                    SUMMARIES
                   %
Result           Query
  No.    Score   Match Length DB  ID                              Description
------------------------------------------------------------------------------
   1       51   100.0       9  21  AAY94202                        Human cytotoxic T
   2       51   100.0       9  21  AAY98670                        WT1 derived immuno
   3       51   100.0       9  21  AAY98778                        WT1 derived immuno
   4       51   100.0       9  21  AAY80200                        Human Wilms' tumou
   5       51   100.0      23  21  AAY98502                        Human WT1 peptide
   6       51   100.0      23  21  AAY98503                        Mouse WT1 peptide
   7       51   100.0     154  19  AAW47175                        Wilms' tumour poly
   8       51   100.0     210  16  AAR85066                        Wilms' tumour anti
   9       51   100.0     210  18  AAW22883                        Wilms' tumour prot
  10       51   100.0     345  12  AAR12240                        Wilm's tumour gene
  11       51   100.0     345  19  AAW47173                        Wilms' tumour poly
```

```
  12       51   100.0     429  16  AAR85065                        Wilms' tumour WT1
  13       51   100.0     449  19  AAW47176                        Wilms' tumour poly
  14       51   100.0     449  21  AAY98804                        Human WT1 protein
  15       51   100.0     449  21  AAY98805                        Mouse WT1 protein
  16       51   100.0     449  21  AAY80196                        Mouse Wilms' tumou
  17       51   100.0     449  21  AAY80197                        Human Wilms' tumou
  18       47    92.2       9  21  AAY98523                        WT1 derived immuno
  19       46    90.2       9  21  AAY98809                        WT1 related peptid
  20       41    80.4       9  21  AAY98564                        WT1 derived immuno
  21       41    80.4       9  21  AAY98752                        WT1 derived immuno
  22       40    78.4       9  21  AAY98808                        WT1 related peptid
  23       37    72.5    1232  17  AAR98217                        Neuronal apoptosis
  24       37    72.5    1295  20  AAY14080                        Gonadotropic hormo
  25       37    72.5    1295  20  AAY09540                        Human apoptosis in
  26       37    72.5    1403  18  AAW20032                        Neuronal apoptosis
  27       37    72.5    1403  18  AAW20033                        Neuronal apoptosis
  28       37    72.5    1403  20  AAY14079                        Gonadotropic hormo
  29       37    72.5    1403  20  AAY09539                        Human apoptosis in
  30       37    72.5    1403  21  AAY88053                        Human NAIP protein
  31       36    70.6     123  21  AAG01172                        Human secreted pro
  32       36    70.6     123  21  AAG01173                        Human secreted pro
  33       36    70.6     200  19  AAY86132                        S. pneumoniae deri
  34       36    70.6     218   9  AAP82925                       Glutathione S-tran
  35       36    70.6     218   9  AAP82926                       Glutathione S-tran
  36       36    70.6     218   9  AAP82927                       Glutathione S-tran
  37       36    70.6     218  13  AAR20031                        Rat liver glutathi
  38       36    70.6     218  13  AAR20032                        Rat liver glutathi
  39       36    70.6     218  13  AAR20033                        Rat brain glutathi
  40       36    70.6     229  21  AAB58846                        Breast and ovarian
  41       36    70.6     369  21  AAG40619                        Arabidopsis thalia
  42       36    70.6     370  20  AAY06623                        Human tyrosylprote
  43       36    70.6     370  20  AAY06624                        Mouse tyrosylprote
  44       36    70.6     370  21  AAY84304                        A human tyrosylpro
  45       36    70.6     370  21  AAY84305                        A murine tyrosylpr


                                 ALIGNMENTS


RESULT   1
AAY94202
ID   AAY94202 standard; peptide; 9 AA.
XX
AC   AAY94202;
XX
DT   28-JUL-2000  (first entry)
XX
DE   Human cytotoxic T lymphocyte-recognised WT1 peptide WT126-34.
XX
KW   WT126-3; peptide; epitope; Wilm's tumour gene; leukaemia;
KW   breast cancer; melanoma; ovarian cancer; immunotherapy.
XX
OS   Homo sapiens.
XX
PN   WO200026249-A1.
XX
PD   11-MAY-2000.
XX
PF   02-NOV-1999;   99WO-GB03572.
XX
PR   02-NOV-1998;   98GB-0023897.
XX
PA   (IMCO-) IMPERIAL COLLEGE INNOVATIONS LTD.
XX
PI   Stauss HJ,  Gao L;
XX
DR   WPI; 2000-376123/32.
XX
PT   Novel peptides comprising WT-1 and GATA-1 epitopes, their fragments or
PT   variants, useful as vaccines for cancer immunotherapy  -
XX
PS   Claim 1; Page 74; 93pp; English.
XX
```

```
CC   The present sequence is peptide epitope WT126-34, produced by WT1
CC   expressing cells and found at residues 126-134 of the WT1 protein, which
CC   is recognised by cytotoxic T lymphocytes. WT1 is aberrantly expressed in
CC   leukaemias, breast cancer, melanoma and ovarian cancer. The peptide can
CC   be used as a vaccine to stimulate the elimination, by cytotoxic T
CC   lymphocytes, of cancer cells aberrantly expressing WT1. In addition, the
CC   nucleic acid encoding the peptide may also be used in the same manner.
CC   Alternatively, the peptide may be used in vitro to produce activated
CC   cytotoxic T lymphocytes.
XX
SQ   Sequence   9 AA;

  Query Match             100.0%;  Score 51;  DB 21;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 3.4e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 RMFPNAPYL 9
            |||||||||
Db        1 rmfpnapyl 9


RESULT   2
AAY98670
ID   AAY98670 standard; Peptide; 9 AA.
XX
AC   AAY98670;
XX
DT   31-JUL-2000  (first entry)
XX
DE   WT1 derived immunogenic peptide SEQ ID NO:185.
XX
KW   WT1; immunotherapy; immunogenic; malignant disease; cancer; leukaemia;
KW   metastatic disease; mouse; human; Wilm's tumour; immune response;
KW   vaccine.
XX
OS   Homo sapiens.
XX
PN   WO200018795-A2.
XX
PD   06-APR-2000.
XX
PF   30-SEP-1999;   99WO-US22819.
XX
PR   30-SEP-1998;   98US-0164223.
PR   25-MAR-1999;   99US-0276484.
XX
PA   (CORI-) CORIXA CORP.
PA   (GAIG/) GAIGER A.
XX
PI   Gaiger A,  Cheever M;
XX
DR   WPI; 2000-293107/25.
XX
PT   Novel polypeptides comprising an immunogenic portion of a native WT1
PT   polypeptide, useful for inhibiting the development of malignant
PT   diseases associated with WT1 expression e.g. leukemia or cancer  -
XX
PS   Claim 4; Page 171; 193pp; English.
XX
CC   The present invention describes polypeptides (I) comprising an
CC   immunogenic portion of a native Wilm's Tumour gene product polypeptide,
CC   WT1, (or variants of the immunogenic portion retaining the ability to
CC   react with WT1-specific antisera and/or T-cell lines or clones) and
CC   comprising 16 consecutive amino acids (aa) or less of a native WT1
CC   polypeptide. The polypeptides are useful therapeutically and to
CC   manufacture medicaments for enhancing/inducing an immune response in
CC   patients. The polypeptides, mimetics or polynucleotides can be included
CC   with a carrier/excipient in pharmaceutical compositions or with a
CC   non-specific immune response enhancer (e.g. an adjuvant or enhancer
CC   preferentially enhancing a T cell response) in vaccines. Pharmaceutical
CC   compositions and vaccines can be administered to human patients to
CC   enhance or induce an immune response specific for WT1 or a cell
CC   expressing WT1, useful to inhibit the development of malignant diseases
CC   associated with WT1 expression, e.g. leukemia (especially acute/chronic
CC   myeloid leukemia or acute lymphocytic leukemia) or cancer (especially
CC   breast, lung, thyroid or gastrointestinal cancer, or a melanoma).
CC   AAY98501 to AAY98811 represent polypeptide sequences, and AAA13848 to
CC   AAA13862 represent PCR primers, used in the exemplification of the
CC   present invention.
XX
SQ   Sequence   9 AA;

  Query Match             100.0%;  Score 51;  DB 21;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 3.4e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 RMFPNAPYL 9
            |||||||||
Db        1 rmfpnapyl 9


RESULT   3
AAY98778
ID   AAY98778 standard; Peptide; 9 AA.
XX
AC   AAY98778;
XX
DT   31-JUL-2000  (first entry)
XX
DE   WT1 derived immunogenic peptide SEQ ID NO:293.
XX
KW   WT1; immunotherapy; immunogenic; malignant disease; cancer; leukaemia;
KW   metastatic disease; mouse; human; Wilm's tumour; immune response;
KW   vaccine.
XX
OS   Mus musculus.
XX
PN   WO200018795-A2.
XX
PD   06-APR-2000.
XX
PF   30-SEP-1999;   99WO-US22819.
XX
PR   30-SEP-1998;   98US-0164223.
PR   25-MAR-1999;   99US-0276484.
XX
PA   (CORI-) CORIXA CORP.
PA   (GAIG/) GAIGER A.
XX
PI   Gaiger A,  Cheever M;
XX
DR   WPI; 2000-293107/25.
XX
PT   Novel polypeptides comprising an immunogenic portion of a native WT1
PT   polypeptide, useful for inhibiting the development of malignant
PT   diseases associated with WT1 expression e.g. leukemia or cancer  -
XX
PS   Claim 4; Page 186; 193pp; English.
XX
CC   The present invention describes polypeptides (I) comprising an
CC   immunogenic portion of a native Wilm's Tumour gene product polypeptide,
CC   WT1, (or variants of the immunogenic portion retaining the ability to
CC   react with WT1-specific antisera and/or T-cell lines or clones) and
CC   comprising 16 consecutive amino acids (aa) or less of a native WT1
CC   polypeptide. The polypeptides are useful therapeutically and to
CC   manufacture medicaments for enhancing/inducing an immune response in
CC   patients. The polypeptides, mimetics or polynucleotides can be included
CC   with a carrier/excipient in pharmaceutical compositions or with a
CC   non-specific immune response enhancer (e.g. an adjuvant or enhancer
CC   preferentially enhancing a T cell response) in vaccines. Pharmaceutical
CC   compositions and vaccines can be administered to human patients to
CC   enhance or induce an immune response specific for WT1 or a cell
CC   expressing WT1, useful to inhibit the development of malignant diseases
CC   associated with WT1 expression, e.g. leukemia (especially acute/chronic
```

```
CC   myeloid leukemia or acute lymphocytic leukemia) or cancer (especially
CC   breast, lung, thyroid or gastrointestinal cancer, or a melanoma).
CC   AAY98501 to AAY98811 represent polypeptide sequences, and AAA13848 to
CC   AAA13862 represent PCR primers, used in the exemplification of the
CC   present invention.
XX
SQ   Sequence   9 AA;


  Query Match             100.0%;  Score 51;  DB 21;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 3.4e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         1 RMFPNAPYL 9
             |||||||||
Db         1 rmfpnapyl 9


RESULT   4
AAY80200
ID   AAY80200 standard; Peptide; 9 AA.
XX
AC   AAY80200;
XX
DT   24-MAY-2000  (first entry)
XX
DE   Human Wilms' tumour suppressor gene WT1 product peptide SEQ ID NO:5.
XX
KW   Wilm's tumour suppressor gene; WT1; cancer; antigen; vaccine; MHC;
KW   major histocompatibility complex; leukaemia; tumour; antitumour.
XX
OS   Homo sapiens.
XX
PN   WO200006602-A1.
XX
PD   10-FEB-2000.
XX
PF   30-JUL-1999;   99WO-JP04130.
XX
PR   31-JUL-1998;   98JP-0218093.
XX
PA   (SUGI/) SUGIYAMA H.
XX
PI   Sugiyama H,  Oka Y;
XX
DR   WPI; 2000-195264/17.
XX
PT   Cancer antigens based on Wilm's tumor suppressor gene WT1 product or
PT   peptide derivatives, for cancer vaccines in treating leukemia and solid
PT   tumors e.g. stomach cancer, skin cancer, lung cancer and breast cancer
XX
PS   Claim 5; Page 18; 48pp; Japanese.
XX
CC   The present invention describes a cancer antigen containing the active
CC   component of Wilms' tumour suppressor gene WT1 product, or partial
CC   peptides, for cancer vaccines in treating leukaemia and solid tumours.
CC   The cancer antigens are useful for cancer vaccines in treating
CC   leukaemia, bone-marrow abnormal formation syndrome, malignant lymphoma,
CC   multiple myeloma, stomach cancer, cancer of the large intestine, lung
CC   cancer, breast cancer, blastoma, liver cancer, skin cancer, bladder
CC   cancer, prostate cancer, uterus cancer, cervical cancer, or ovary
CC   cancer. The present sequence represents a peptide from the human
CC   Wilms' tumour suppressor gene WT1 product.
XX
SQ   Sequence   9 AA;


  Query Match             100.0%;  Score 51;  DB 21;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 3.4e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         1 RMFPNAPYL 9
             |||||||||
```

```
Db         1 rmfpnapyl 9


RESULT   5
AAY98502
ID   AAY98502 standard; Peptide; 23 AA.
XX
AC   AAY98502;
XX
DT   31-JUL-2000  (first entry)
XX
DE   Human WT1 peptide SEQ ID NO:2.
XX
KW   WT1; immunotherapy; immunogenic; malignant disease; cancer; leukaemia;
KW   metastatic disease; mouse; human; Wilm's tumour; immune response;
KW   vaccine.
XX
OS   Homo sapiens.
XX
PN   WO200018795-A2.
XX
PD   06-APR-2000.
XX
PF   30-SEP-1999;   99WO-US22819.
XX
PR   30-SEP-1998;   98US-0164223.
PR   25-MAR-1999;   99US-0276484.
XX
PA   (CORI-) CORIXA CORP.
PA   (GAIG/) GAIGER A.
XX
PI   Gaiger A,  Cheever M;
XX
DR   WPI; 2000-293107/25.
XX
PT   Novel polypeptides comprising an immunogenic portion of a native WT1
PT   polypeptide, useful for inhibiting the development of malignant
PT   diseases associated with WT1 expression e.g. leukemia or cancer  -
XX
PS   Claim 4; Page 46; 193pp; English.
XX
CC   The present invention describes polypeptides (I) comprising an
CC   immunogenic portion of a native Wilm's Tumour gene product polypeptide,
CC   WT1, (or variants of the immunogenic portion retaining the ability to
CC   react with WT1-specific antisera and/or T-cell lines or clones) and
CC   comprising 16 consecutive amino acids (aa) or less of a native WT1
CC   polypeptide. The polypeptides are useful therapeutically and to
CC   manufacture medicaments for enhancing/inducing an immune response in
CC   patients. The polypeptides, mimetics or polynucleotides can be included
CC   with a carrier/excipient in pharmaceutical compositions or with a
CC   non-specific immune response enhancer (e.g. an adjuvant or enhancer
CC   preferentially enhancing a T cell response) in vaccines. Pharmaceutical
CC   compositions and vaccines can be administered to human patients to
CC   enhance or induce an immune response specific for WT1 or a cell
CC   expressing WT1, useful to inhibit the development of malignant diseases
CC   associated with WT1 expression, e.g. leukemia (especially acute/chronic
CC   myeloid leukemia or acute lymphocytic leukemia) or cancer (especially
CC   breast, lung, thyroid or gastrointestinal cancer, or a melanoma).
CC   AAY98501 to AAY98811 represent polypeptide sequences, and AAA13848 to
CC   AAA13862 represent PCR primers, used in the exemplification of the
CC   present invention.
XX
SQ   Sequence   23 AA;


  Query Match             100.0%;  Score 51;  DB 21;  Length 23;
  Best Local Similarity   100.0%;  Pred. No. 0.0027;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         1 RMFPNAPYL 9
             |||||||||
Db        10 rmfpnapyl 18
```

```
RESULT   6
AAY98503
ID   AAY98503 standard; Peptide; 23 AA.
XX
AC   AAY98503;
XX
DT   31-JUL-2000  (first entry)
XX
DE   Mouse WT1 peptide SEQ ID NO:3.
XX
KW   WT1; immunotherapy; immunogenic; malignant disease; cancer; leukaemia;
KW   metastatic disease; mouse; human; Wilm's tumour; immune response;
KW   vaccine.
XX
OS   Mus musculus.
XX
PN   WO200018795-A2.
XX
PD   06-APR-2000.
XX
PF   30-SEP-1999;   99WO-US22819.
XX
PR   30-SEP-1998;   98US-0164223.
PR   25-MAR-1999;   99US-0276484.
XX
PA   (CORI-) CORIXA CORP.
PA   (GAIG/) GAIGER A.
XX
PI   Gaiger A,  Cheever M;
XX
DR   WPI; 2000-293107/25.
XX
PT   Novel polypeptides comprising an immunogenic portion of a native WT1
PT   polypeptide, useful for inhibiting the development of malignant
PT   diseases associated with WT1 expression e.g. leukemia or cancer  -
XX
PS   Example 3; Page 46; 193pp; English.
XX
CC   The present invention describes polypeptides (I) comprising an
CC   immunogenic portion of a native Wilm's Tumour gene product polypeptide,
CC   WT1, (or variants of the immunogenic portion retaining the ability to
CC   react with WT1-specific antisera and/or T-cell lines or clones) and
CC   comprising 16 consecutive amino acids (aa) or less of a native WT1
CC   polypeptide. The polypeptides are useful therapeutically and to
CC   manufacture medicaments for enhancing/inducing an immune response in
CC   patients. The polypeptides, mimetics or polynucleotides can be included
CC   with a carrier/excipient in pharmaceutical compositions or with a
CC   non-specific immune response enhancer (e.g. an adjuvant or enhancer
CC   preferentially enhancing a T cell response) in vaccines. Pharmaceutical
CC   compositions and vaccines can be administered to human patients to
CC   enhance or induce an immune response specific for WT1 or a cell
CC   expressing WT1, useful to inhibit the development of malignant diseases
CC   associated with WT1 expression, e.g. leukemia (especially acute/chronic
CC   myeloid leukemia or acute lymphocytic leukemia) or cancer (especially
CC   breast, lung, thyroid or gastrointestinal cancer, or a melanoma).
CC   AAY98501 to AAY98811 represent polypeptide sequences, and AAA13848 to
CC   AAA13862 represent PCR primers, used in the exemplification of the
CC   present invention.
XX
SQ   Sequence   23 AA;

  Query Match             100.0%;  Score 51;  DB 21;  Length 23;
  Best Local Similarity   100.0%;  Pred. No. 0.0027;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 RMFPNAPYL 9
          |||||||||
Db     10 rmfpnapyl 18

RESULT   7
AAW47175
ID   AAW47175 standard; Protein; 154 AA.
XX
AC   AAW47175;
XX
DT   01-JUN-1998  (first entry)
XX
DE   Wilms' tumour polypeptide (WT33) proline and glutamine rich region.
XX
KW   Wilms' tumour; WT33; cancer treatment; antibody production; WAGR;
KW   Denys-drash syndrome; WT1.
XX
OS   Homo sapiens.
XX
PN   US5726288-A.
XX
PD   10-MAR-1998.
XX
PF   02-AUG-1993;   93US-0102942.
XX
PR   02-AUG-1993;   93US-0102942.
PR   13-NOV-1989;   89US-0435780.
PR   13-NOV-1990;   90US-0614161.
XX
PA   (MASI ) MASSACHUSETTS INST TECHNOLOGY.
XX
PI   Bruening W,  Buckler AJ,  Call KM,  Darveau A,  Glaser TM;
PI   Haber DA,  Housman DE,  Ito CY,  Pelletier J,  Rose EA;
XX
DR   WPI; 1998-192828/17.
XX
PT   New Wilms' tumour polypeptide - useful for, e.g. cancer treatment
PT   and antibody production
XX
PS   Claim 2; Fig 4; 42pp; English.
XX
CC   This sequence represents the proline and glutamine rich region of a new
CC   Wilms' tumour (WT33) polypeptide. The Wilms' tumour gene (WT1) is
CC   associated with 11p3 locus on the human chromosome. This proline and
CC   glutamine rich region indicates that the Wilms' tumour polypeptide has
CC   a role in transcription regulation. The polypeptide or immunogenic
CC   fragments of the polypeptide can be used to treat cancerous or
CC   precancerous conditions (especially Wilms' tumour), or to raise
CC   antibodies for diagnostic use. The product allow detection of risk of
CC   development of Wilms' tumour, e.g. in diseases such as WAGR and
CC   Denys-Drash syndrome, to be assessed prior to current methods of
CC   detection.
XX
SQ   Sequence   154 AA;

  Query Match             100.0%;  Score 51;  DB 19;  Length 154;
  Best Local Similarity   100.0%;  Pred. No. 0.021;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 RMFPNAPYL 9
          |||||||||
Db     40 rmfpnapyl 48

RESULT   8
AAR85066
ID   AAR85066 standard; Protein; 210 AA.
XX
AC   AAR85066;
XX
DT   02-FEB-1996  (first entry)
XX
DE   Wilms' tumour antigen WT1-6F.
XX
KW   Wilms' tumour; WT1-6F; monoclonal antibody; leukaemia.
XX
OS   Chimeric Homo sapiens;.
```

```
OS   Chimeric synthetic.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1..11
FT                   /label= Histidine_fusion_peptide
FT                   /note= "hexa-histidine peptide used to facilitate
FT                          fusion protein purification"
FT   Protein         12..192
FT                   /label= WT1
FT   Peptide         193..219
FT                   /note= "vector-derived amino acids"
XX
PN   WO9529995-A1.
XX
PD   09-NOV-1995.
XX
PF   25-APR-1995;   95WO-US05523.
XX
PR   28-APR-1994;   94US-0234783.
XX
PA   (WIST-) WISTAR INST ANATOMY & BIOLOGY.
XX
PI   Herlyn M,  Morris J,  Rauscher FJ,  Rodeck U;
XX
DR   WPI; 1995-393072/50.
DR   N-PSDB; AAT02462.
XX
PT   Monoclonal antibodies specific for Wilms' tumour protein antigen WT1
PT   - useful for detecting, monitoring and diagnosing malignancies
PT   characterised by expression of WT1 protein, e.g. leukaemia
XX
PS   Example 1; Page 37-38; 54pp; English.
XX
CC   Wilms' tumor antigen WT1-6F (AAR85066) is based on amino acids
CC   1-181 of the native sequence, plus additional N- and C-terminal
CC   sequences, and is encoded by synthetic DNA (AAT02462) optimized for
CC   expression in E. coli.  WT1-6F is used to raise monoclonal
CC   antibodies specific for WT1.
XX
SQ   Sequence   210 AA;


  Query Match             100.0%;  Score 51;  DB 16;  Length 210;
  Best Local Similarity   100.0%;  Pred. No. 0.03;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RMFPNAPYL 9
             |||||||||
Db     137 rmfpnapyl 145


RESULT   9
AAW22883
ID   AAW22883 standard; Protein; 210 AA.
XX
AC   AAW22883;
XX
DT   24-SEP-1997  (first entry)
XX
DE   Wilms' tumour protein antigen WT1-6F.
XX
KW   Wilms' tumour; WT1; antigen; WT1-6F; monoclonal; antibody;
KW   diagnosis; mesothelioma; prostate; ovarian; cancer; leukaemia;
KW   leukemia.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1..11
FT                   /note= "histidine fusion peptide to facilitate
FT                          purification"
FT   Protein         12..192
```

```
FT                   /note= "amino acids 1-181 of WT1 protein"
FT   Region          193..210
FT                   /note= "vector sequences added during cloning"
XX
PN   US5633142-A.
XX
PD   27-MAY-1997.
XX
PF   28-APR-1994;   94US-0234783.
XX
PR   01-JUN-1995;   95US-0456907.
PR   28-APR-1994;   94US-0234783.
XX
PA   (WIST-) WISTAR INST ANATOMY & BIOLOGY.
XX
PI   Herlyn M,  Morris J,  Rauscher FJ,  Rodeck U;
XX
DR   WPI; 1997-297358/27.
DR   N-PSDB; AAT75529.
XX
PT   Diagnosis of mesothelioma and monitoring of leukaemia therapy -
PT   using monoclonal antibodies against Wilms' tumour 1 antigen
XX
PS   Example 1; Columns 17-20; 21pp; English.
XX
CC   The present sequence is the Wilms' tumour protein antigen
CC   WT1-6F, which comprises amino acids 1-181 of the Wilms' tumour 1
CC   antigen (WT1) protein. WT1-6F can be used to raise monoclonal
CC   antibodies (MAb), e.g. H2, H7 and/or H17, which are secreted by
CC   the hybridoma cell lines ATCC 11598, 11599 and 11600. The MAb can
CC   be used to diagnose mesothelioma, prostate cancer, ovarian cancer
CC   or leukaemia by binding an antigen in a whole blood, serum, plasma,
CC   synovial fluid or tissue sample, or monitor therapy in leukaemia
CC   patients by binding an antigen in a whole blood, plasma, serum,
CC   urine or bone marrow sample, indicating the presence of active
CC   leukaemia cells.
XX
SQ   Sequence   210 AA;

  Query Match             100.0%;  Score 51;  DB 18;  Length 210;
  Best Local Similarity   100.0%;  Pred. No. 0.03;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RMFPNAPYL 9
             |||||||||
Db     137 rmfpnapyl 145


RESULT  10
AAR12240
ID   AAR12240 standard; Protein; 345 AA.
XX
AC   AAR12240;
XX
DT   15-AUG-1991  (first entry)
XX
DE   Wilm's tumour gene prod.
XX
KW   Wilm's tumour; monoclonal antibodies.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Domain          3..156
FT                   /label= proline/glutamine rich region
FT   Domain          216..226
FT                   /label= zinc finger domain 1
FT   Domain          227..246
FT                   /label= zinc finger domain 2
FT   Domain          247..276
FT                   /label= zinc finger domain 3
FT   Domain          277..304
```

```
FT                    /label= zinc finger domain 4
XX
PN   WO9107509-A.
XX
PD   30-MAY-1991.
XX
PF   13-NOV-1990;   90WO-US06629.
XX
PR   13-NOV-1989;   89US-0435780.
XX
PA   (MASI ) MASSACHUSETTS INST TECH.
XX
PI   Call KM,  Glaser TM,  Ito CY,  Buckler AJ,  Pelletier J;
PI   Haber DA,  Rose EA,  Housman DE;
XX
DR   WPI; 1991-178117/24.
DR   N-PSDB; AAQ12020.
XX
PT   Nucleotide sequence of human chromosome 11 band 13 Wilm's tumour
PT   locus - used to diagnose, quantify and treat Wilm's tumours
XX
PS   Disclosure; fig 3; 67pp; English.
XX
CC   This polypeptide is encoded by the Wilm's tumour (WT) gene which
CC   maps to the 11p13 locus of human chromosome 11.  It contains 4
CC   zinc finger binding domains and a proline/glutamine rich region
CC   and hence is thought to be involved in transcription regulation.
CC   Using this WT protein and the DNA, probes and antibodies can be
CC   developed for diagnosing WT.  Lesions similar to WT may also
CC   be detected.
XX
SQ   Sequence   345 AA;


  Query Match             100.0%;  Score 51;  DB 12;  Length 345;
  Best Local Similarity   100.0%;  Pred. No. 0.051;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RMFPNAPYL 9
             |||||||||
Db      42 rmfpnapyl 50


RESULT  11
AAW47173
ID   AAW47173 standard; Protein; 345 AA.
XX
AC   AAW47173;
XX
DT   01-JUN-1998  (first entry)
XX
DE   Wilms' tumour polypeptide (WT33).
XX
KW   Wilms' tumour; WT33; cancer treatment; antibody production; WAGR;
KW   Denys-drash syndrome; WT1.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Region          3..156
FT                   /note= "proline and glutamine rich region"
XX
PN   US5726288-A.
XX
PD   10-MAR-1998.
XX
PF   02-AUG-1993;   93US-0102942.
XX
PR   02-AUG-1993;   93US-0102942.
PR   13-NOV-1989;   89US-0435780.
PR   13-NOV-1990;   90US-0614161.
XX
PA   (MASI ) MASSACHUSETTS INST TECHNOLOGY.
```

```
XX
PI   Bruening W,  Buckler AJ,  Call KM,  Darveau A,  Glaser TM;
PI   Haber DA,  Housman DE,  Ito CY,  Pelletier J,  Rose EA;
XX
DR   WPI; 1998-192828/17.
DR   N-PSDB; AAV17060.
XX
PT   New Wilms' tumour polypeptide - useful for, e.g. cancer treatment
PT   and antibody production
XX
PS   Claim 2; Fig 3; 42pp; English.
XX
CC   This is a new Wilms' tumour (WT33) polypeptide. The Wilms' tumour gene
CC   (WT1) is associated with 11p3 locus on the human chromosome. This
CC   polypeptide has a region rich in proline and glutamine (AAW47175)
CC   indicating that it has a role in transcription regulation. The
CC   polypeptide or immunogenic fragments of the polypeptide can be used to
CC   treat cancerous or precancerous conditions (especially Wilms' tumour),
CC   or to raise antibodies for diagnostic use. The product allow detection
CC   of risk of development of Wilms' tumour, e.g. in diseases such as WAGR
CC   and Denys-Drash syndrome, to be assessed prior to current methods of
CC   detection.
XX
SQ   Sequence   345 AA;


  Query Match             100.0%;  Score 51;  DB 19;  Length 345;
  Best Local Similarity   100.0%;  Pred. No. 0.051;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RMFPNAPYL 9
             |||||||||
Db      42 rmfpnapyl 50


RESULT  12
AAR85065
ID   AAR85065 standard; Protein; 429 AA.
XX
AC   AAR85065;
XX
DT   02-FEB-1996  (first entry)
XX
DE   Wilms' tumour WT1 antigen.
XX
KW   Wilms' tumour; WT1; monoclonal antibody; leukaemia.
XX
OS   Homo sapiens.
XX
PN   WO9529995-A1.
XX
PD   09-NOV-1995.
XX
PF   25-APR-1995;   95WO-US05523.
XX
PR   28-APR-1994;   94US-0234783.
XX
PA   (WIST-) WISTAR INST ANATOMY & BIOLOGY.
XX
PI   Herlyn M,  Morris J,  Rauscher FJ,  Rodeck U;
XX
DR   WPI; 1995-393072/50.
DR   N-PSDB; AAR85065.
XX
PT   Monoclonal antibodies specific for Wilms' tumour protein antigen WT1
PT   - useful for detecting, monitoring and diagnosing malignancies
PT   characterised by expression of WT1 protein, e.g. leukaemia
XX
PS   Claim 1; Page 42-43; 54pp; English.
XX
CC   The human Wilm's tumour antigen WT1 (sequence given in AAR85065) was
CC   used as the basis for the design of a synthetic WT1 sequence (see
CC   AAR85066) used to raise monoclonal antibodies specific for the antigen.
```

```
XX
SQ   Sequence   429 AA;

  Query Match             100.0%;  Score 51;  DB 16;  Length 429;
  Best Local Similarity   100.0%;  Pred. No. 0.064;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RMFPNAPYL 9
           |||||||||
Db     126 rmfpnapyl 134


RESULT  13
AAW47176
ID   AAW47176 standard; Protein; 449 AA.
XX
AC   AAW47176;
XX
DT   01-JUN-1998  (first entry)
XX
DE   Wilms' tumour polypeptide (WT1).
XX
KW   Wilms' tumour; WT33; cancer treatment; antibody production; WAGR;
KW   Denys-drash syndrome; WT1.
XX
OS   Homo sapiens.
XX
PN   US5726288-A.
XX
PD   10-MAR-1998.
XX
PF   02-AUG-1993;   93US-0102942.
XX
PR   02-AUG-1993;   93US-0102942.
PR   13-NOV-1989;   89US-0435780.
PR   13-NOV-1990;   90US-0614161.
XX
PA   (MASI ) MASSACHUSETTS INST TECHNOLOGY.
XX
PI   Bruening W,  Buckler AJ,  Call KM,  Darveau A,  Glaser TM;
PI   Haber DA,  Housman DE,  Ito CY,  Pelletier J,  Rose EA;
XX
DR   WPI; 1998-192828/17.
XX
PT   New Wilms' tumour polypeptide - useful for, e.g. cancer treatment
PT   and antibody production
XX
PS   Disclosure; Columns 39-42; 42pp; English.
XX
CC   This is a Wilms' tumour (WT1) polypeptide. The Wilms' tumour gene
CC   (WT1) is associated with 11p3 locus on the human chromosome. The
CC   polypeptide or immunogenic fragments of the polypeptide can be used to
CC   treat cancerous or precancerous conditions (especially Wilms' tumour), or
CC   to raise antibodies for diagnostic use. The product allow detection of
CC   risk of development of Wilms' tumour, e.g. in diseases such as WAGR and
CC   Denys-Drash syndrome; to be assessed prior to current methods of
CC   detection.
XX
SQ   Sequence   449 AA;

  Query Match             100.0%;  Score 51;  DB 19;  Length 449;
  Best Local Similarity   100.0%;  Pred. No. 0.067;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RMFPNAPYL 9
           |||||||||
Db     126 rmfpnapyl 134


RESULT  14
AAY98804
ID   AAY98804 standard; Protein; 449 AA.
XX
AC   AAY98804;
XX
DT   31-JUL-2000  (first entry)
XX
DE   Human WT1 protein sequence SEQ ID NO:319.
XX
KW   WT1; immunotherapy; immunogenic; malignant disease; cancer; leukaemia;
KW   metastatic disease; mouse; human; Wilm's tumour; immune response;
KW   vaccine.
XX
OS   Homo sapiens.
XX
PN   WO200018795-A2.
XX
PD   06-APR-2000.
XX
PF   30-SEP-1999;   99WO-US22819.
XX
PR   30-SEP-1998;   98US-0164223.
PR   25-MAR-1999;   99US-0276484.
XX
PA   (CORI-) CORIXA CORP.
PA   (GAIG/) GAIGER A.
XX
PI   Gaiger A,  Cheever M;
XX
DR   WPI; 2000-293107/25.
XX
PT   Novel polypeptides comprising an immunogenic portion of a native WT1
PT   polypeptide, useful for inhibiting the development of malignant
PT   diseases associated with WT1 expression e.g. leukemia or cancer  -
XX
PS   Disclosure; Page 190-191; 193pp; English.
XX
CC   The present invention describes polypeptides (I) comprising an
CC   immunogenic portion of a native Wilm's Tumour gene product polypeptide,
CC   WT1, (or variants of the immunogenic portion retaining the ability to
CC   react with WT1-specific antisera and/or T-cell lines or clones) and
CC   comprising 16 consecutive amino acids (aa) or less of a native WT1
CC   polypeptide. The polypeptides are useful therapeutically and to
CC   manufacture medicaments for enhancing/inducing an immune response in
CC   patients. The polypeptides, mimetics or polynucleotides can be included
CC   with a carrier/excipient in pharmaceutical compositions or with a
CC   non-specific immune response enhancer (e.g. an adjuvant or enhancer
CC   preferentially enhancing a T cell response) in vaccines. Pharmaceutical
CC   compositions and vaccines can be administered to human patients to
CC   enhance or induce an immune response specific for WT1 or a cell
CC   expressing WT1, useful to inhibit the development of malignant diseases
CC   associated with WT1 expression, e.g. leukemia (especially acute/chronic
CC   myeloid leukemia or acute lymphocytic leukemia) or cancer (especially
CC   breast, lung, thyroid or gastrointestinal cancer, or a melanoma).
CC   AAY98501 to AAY98811 represent polypeptide sequences, and AAA13848 to
CC   AAA13862 represent PCR primers, used in the exemplification of the
CC   present invention.
XX
SQ   Sequence   449 AA;

  Query Match             100.0%;  Score 51;  DB 21;  Length 449;
  Best Local Similarity   100.0%;  Pred. No. 0.067;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RMFPNAPYL 9
           |||||||||
Db     126 rmfpnapyl 134


RESULT  15
AAY98805
ID   AAY98805 standard; Protein; 449 AA.
XX
```

```
AC   AAY98805;
XX
DT   31-JUL-2000  (first entry)
XX
DE   Mouse WT1 protein sequence SEQ ID NO:320.
XX
KW   WT1; immunotherapy; immunogenic; malignant disease; cancer; leukaemia;
KW   metastatic disease; mouse; human; Wilm's tumour; immune response;
KW   vaccine.
XX
OS   Mus musculus.
XX
PN   WO200018795-A2.
XX
PD   06-APR-2000.
XX
PF   30-SEP-1999;   99WO-US22819.
XX
PR   30-SEP-1998;   98US-0164223.
PR   25-MAR-1999;   99US-0276484.
XX
PA   (CORI-) CORIXA CORP.
PA   (GAIG/) GAIGER A.
XX
PI   Gaiger A,  Cheever M;
XX
DR   WPI; 2000-293107/25.
XX
PT   Novel polypeptides comprising an immunogenic portion of a native WT1
PT   polypeptide, useful for inhibiting the development of malignant
PT   diseases associated with WT1 expression e.g. leukemia or cancer  -
XX
PS   Disclosure; Page 191-192; 193pp; English.
XX
CC   The present invention describes polypeptides (I) comprising an
CC   immunogenic portion of a native Wilm's Tumour gene product polypeptide,
CC   WT1, (or variants of the immunogenic portion retaining the ability to
CC   react with WT1-specific antisera and/or T-cell lines or clones) and
CC   comprising 16 consecutive amino acids (aa) or less of a native WT1
CC   polypeptide. The polypeptides are useful therapeutically and to
CC   manufacture medicaments for enhancing/inducing an immune response in
CC   patients. The polypeptides, mimetics or polynucleotides can be included
CC   with a carrier/excipient in pharmaceutical compositions or with a
CC   non-specific immune response enhancer (e.g. an adjuvant or enhancer
CC   preferentially enhancing a T cell response) in vaccines. Pharmaceutical
CC   compositions and vaccines can be administered to human patients to
CC   enhance or induce an immune response specific for WT1 or a cell
CC   expressing WT1, useful to inhibit the development of malignant diseases
CC   associated with WT1 expression, e.g. leukemia (especially acute/chronic
CC   myeloid leukemia or acute lymphocytic leukemia) or cancer (especially
CC   breast, lung, thyroid or gastrointestinal cancer, or a melanoma).
CC   AAY98501 to AAY98811 represent polypeptide sequences, and AAA13848 to
CC   AAA13862 represent PCR primers, used in the exemplification of the
CC   present invention.
XX
SQ   Sequence   449 AA;

  Query Match             100.0%;  Score 51;  DB 21;  Length 449;
  Best Local Similarity   100.0%;  Pred. No. 0.067;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RMFPNAPYL 9
             |||||||||
Db     126 rmfpnapyl 134


Search completed: October 23, 2001, 13:21:01
Job time: 54 sec
```

```
                       GenCore version 4.5
             Copyright (c) 1993 - 2000  Compugen Ltd.

OM protein - protein search, using sw model

Run on:          October 23, 2001, 13:21:07 ; Search time 47.11 Seconds
                                              (without alignments)
                                              25.276 Million cell updates/sec

Title:           US-09-625-963-1
Perfect score:   51
Sequence:        1 RMFPNAPYL 9

Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched:        425026 seqs, 132305027 residues

Total number of hits satisfying chosen parameters:      425026

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       SPTREMBL_16:*
                1:  sp_archea:*
                2:  sp_bacteria:*
                3:  sp_fungi:*
                4:  sp_human:*
                5:  sp_invertebrate:*
                6:  sp_mammal:*
                7:  sp_mhc:*
                8:  sp_organelle:*
                9:  sp_phage:*
                10:  sp_plant:*
                11:  sp_rodent:*
                12:  sp_unclassified:*
                13:  sp_vertebrate:*
                14:  sp_virus:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                                   SUMMARIES
                     %
Result               Query
  No.     Score      Match Length DB  ID                              Description
-------------------------------------------------------------------------------
    1       51      100.0    407  13  P79958                      P79958 xenopus lae
    2       51      100.0    446  4   Q15881                     Q15881 homo sapien
    3       47       92.2    390  13  O42223                      O42223 trachemys s
    4       43       84.3    409  13  Q91657                      Q91657 xenopus lae
    5       40       78.4    392  13  Q91030                      Q91030 gallus gall
    6       40       78.4    414  13  Q9I8A1                      Q9i8a1 gallus gall
    7       40       78.4    417  13  Q9I8A0                      Q9i8a0 gallus gall
    8       40       78.4    419  13  Q9PUT7                      Q9put7 brachydanio
    9       39       76.5    203  2   O83076                     O83076 treponema p
   10       39       76.5    392  13  Q9IBF0                      Q9ibf0 anguilla ja
   11       39       76.5    426  13  Q9W611                      Q9w611 cynops pyrr
   12       38       74.5    205  2   Q9I4D6                     Q9i4d6 pseudomonas
   13       38       74.5    581  10  Q9FZ97                      Q9fz97 arabidopsis
   14       38       74.5    662  5   Q9N6S5                     Q9n6s5 leishmania
   15       38       74.5   2175  5   Q9W198                     Q9w198 drosophila
   16       37       72.5    444  10  Q9LS41                      Q9ls41 arabidopsis
   17       36       70.6     32  11  Q9QVM0                      Q9qvm0 mus sp. glu
   18       36       70.6     37  11  Q9QVM3                      Q9qvm3 mus sp. glu
   19       36       70.6     82  4   Q9UE37                     Q9ue37 homo sapien
   20       36       70.6    188  6   Q9MZB4                     Q9mzb4 capra hircu
   21       36       70.6    195  4   Q05465                     Q05465 homo sapien
   22       36       70.6    218  6   Q9TSM5                     Q9tsm5 macaca fasc
   23       36       70.6    218  6   Q9TSM4                     Q9tsm4 macaca fasc
   24       36       70.6    218  6   Q9N0V4                     Q9n0v4 bos taurus
   25       36       70.6    218  11  O35660                      O35660 mus musculu
   26       36       70.6    225  4   O60550                     O60550 homo sapien
   27       36       70.6    225  11  Q9Z1B2                      Q9z1b2 rattus norv
   28       36       70.6    348  8   Q33154                     Q33154 suksdorfia
   29       36       70.6    432  10  Q9SX79                      Q9sx79 arabidopsis
   30       36       70.6    434  10  Q9SX80                      Q9sx80 arabidopsis
   31       36       70.6   1249  11  Q9Z2Q1                      Q9z2q1 rattus norv
   32       35       68.6    149  5   O76502                     O76502 homarus ame
   33       35       68.6    149  5   O76503                     O76503 haliotis ru
   34       35       68.6    150  5   O76504                     O76504 penaeus set
   35       35       68.6    194  2   Q9PGH2                     Q9pgh2 xylella fas
   36       35       68.6    296  2   Q9RKD6                     Q9rkd6 streptomyce
   37       35       68.6    345  2   Q9KVC3                     Q9kvc3 vibrio chol
   38       35       68.6    511  5   Q9GPS2                     Q9gps2 dictyosteli
   39       35       68.6    853  13  O93245                      O93245 oncorhynchu
   40       35       68.6   1072  5   Q9N5Y6                     Q9n5y6 caenorhabdi
   41       35       68.6   1281  14  Q85425                      Q85425 rat cytomeg
   42       34       66.7    124  5   Q9NGQ5                     Q9ngq5 diabrotica
   43       34       66.7    180  2   O67664                     O67664 aquifex aeo
   44       34       66.7    180  10  Q9SVQ5                      Q9svq5 arabidopsis
   45       34       66.7    230  2   Q9L995                     Q9l995 vibrio chol

                               ALIGNMENTS


RESULT   1
P79958
ID   P79958      PRELIMINARY;      PRT;   407 AA.
AC   P79958;
DT   01-MAY-1997 (TrEMBLrel. 03, Created)
DT   01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   WT1 PROTEIN.
GN   WT1.
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=TESTIS;
RX   MEDLINE=97074667; PubMed=8917094;
RA   Semba K., Saito-Ueno R., Takayama G., Kondo M.;
RT   "cDNA cloning and its pronephros-specific expression of the Wilms'
RT   tumor suppressor gene, WT1, from Xenopus laevis.";
RL   Gene 175:167-172(1996).
CC   -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.
DR   EMBL; D82051; BAA11522.1; -.
DR   HSSP; P08046; 1AAY.
DR   InterPro; IPR000822; -.
DR   InterPro; IPR000976; -.
DR   Pfam; PF00096; zf-C2H2; 4.
DR   Pfam; PF02165; WT1; 3.
DR   PRINTS; PR00048; ZINCFINGER.
DR   PROSITE; PS00028; ZINC_FINGER_C2H2; 4.
DR   SMART; SM00355; ZnF_C2H2; 1.
KW   DNA-binding; Metal-binding; Nuclear protein; Zinc-finger.
SQ   SEQUENCE   407 AA;  45983 MW;  E2554C658005870C CRC64;


  Query Match             100.0%;  Score 51;  DB 13;  Length 407;
  Best Local Similarity   100.0%;  Pred. No. 0.084;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RMFPNAPYL 9
```

```
               |||||||||
Db     107 RMFPNAPYL 115


RESULT   2
Q15881
ID   Q15881      PRELIMINARY;      PRT;   446 AA.
AC   Q15881;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   WILMS TUMOR PROTEIN.
GN   WT1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92241883; PubMed=1572653;
RA   Gessler M., Konig A., Bruns G.A.;
RT   "The genomic organization and expression of the WT1 gene.";
RL   Genomics 12:807-813(1992).
CC   -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.
DR   EMBL; X61631; CAA43819.1; -.
DR   EMBL; X61632; CAA43819.1; JOINED.
DR   EMBL; X61633; CAA43819.1; JOINED.
DR   EMBL; X61634; CAA43819.1; JOINED.
DR   EMBL; X61635; CAA43819.1; JOINED.
DR   EMBL; X61636; CAA43819.1; JOINED.
DR   EMBL; X61637; CAA43819.1; JOINED.
DR   EMBL; X61638; CAA43819.1; JOINED.
DR   HSSP; P08046; 1AAY.
DR   InterPro; IPR000822; -.
DR   InterPro; IPR000976; -.
DR   Pfam; PF00096; zf-C2H2; 4.
DR   Pfam; PF02165; WT1; 1.
DR   PRINTS; PR00049; WILMSTUMOUR.
DR   PRINTS; PR00048; ZINCFINGER.
DR   PROSITE; PS00028; ZINC_FINGER_C2H2; 4.
DR   SMART; SM00355; ZnF_C2H2; 1.
KW   DNA-binding; Metal-binding; Nuclear protein; Zinc-finger.
SQ   SEQUENCE   446 AA;  48842 MW;  8CE7FC047F41CF11 CRC64;


  Query Match             100.0%;  Score 51;  DB 4;  Length 446;
  Best Local Similarity   100.0%;  Pred. No. 0.092;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RMFPNAPYL 9
             |||||||||
Db     126 RMFPNAPYL 134


RESULT   3
O42223
ID   O42223      PRELIMINARY;      PRT;   390 AA.
AC   O42223;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   WILMS TUMOR 1 PROTEIN.
GN   TWT1.
OS   Trachemys scripta (Red-eared slider turtle) (Pseudemys scripta).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Testudines; Cryptodira; Testudinoidea; Emydidae; Trachemys.
OX   NCBI_TaxID=34903;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Spotila L.D., Hall S.E.;
RL   Submitted (AUG-1997) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.
DR   EMBL; AF019779; AAB70832.1; -.
DR   HSSP; P08046; 1AAY.
DR   InterPro; IPR000822; -.
DR   InterPro; IPR000976; -.
DR   Pfam; PF00096; zf-C2H2; 4.
DR   Pfam; PF02165; WT1; 2.
DR   PRINTS; PR00048; ZINCFINGER.
DR   PROSITE; PS00028; ZINC_FINGER_C2H2; 3.
DR   SMART; SM00355; ZnF_C2H2; 1.
KW   DNA-binding; Metal-binding; Nuclear protein; Zinc-finger.
SQ   SEQUENCE   390 AA;  43620 MW;  1C9987435BE927C2 CRC64;

  Query Match             92.2%;  Score 47;  DB 13;  Length 390;
  Best Local Similarity   88.9%;  Pred. No. 0.45;
  Matches    8;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 RMFPNAPYL 9
             ||||| |||
Db     109 RMFPNGPYL 117


RESULT   4
Q91657
ID   Q91657      PRELIMINARY;      PRT;   409 AA.
AC   Q91657;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   WILMS' TUMOR SUPPRESSOR (WT1).
GN   WT1.
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Carroll T.J., Vize P.D.;
RL   Submitted (JAN-1996) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE OF 289-370 FROM N.A.
RC   TISSUE=MESONEPHROS;
RX   MEDLINE=96068905; PubMed=7478606;
RA   Kent J., Coriat A.M., Sharpe P.T., Hastie N., Van Heyningen V.;
RT   "The evolution of WT1 sequence and expression pattern in the
RT   vertebrates.";
RL   Oncogene 11:1781-1792(1995).
CC   -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.
DR   EMBL; U42011; AAB53152.1; -.
DR   EMBL; X85733; CAA59738.1; -.
DR   HSSP; P08046; 1AAY.
DR   InterPro; IPR000822; -.
DR   InterPro; IPR000976; -.
DR   Pfam; PF00096; zf-C2H2; 4.
DR   Pfam; PF02165; WT1; 3.
DR   PRINTS; PR00048; ZINCFINGER.
DR   PROSITE; PS00028; ZINC_FINGER_C2H2; 4.
DR   SMART; SM00355; ZnF_C2H2; 1.
KW   DNA-binding; Metal-binding; Nuclear protein; Zinc-finger.
SQ   SEQUENCE   409 AA;  46186 MW;  2217FC04612CDE10 CRC64;


  Query Match             84.3%;  Score 43;  DB 13;  Length 409;
  Best Local Similarity   88.9%;  Pred. No. 2.7;
  Matches    8;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 RMFPNAPYL 9
             ||| |||||
Db     106 RMFSNAPYL 114
```

```
RESULT   5
Q91030
ID   Q91030      PRELIMINARY;      PRT;   392 AA.
AC   Q91030;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   CHICK WILM'S TUMOUR PROTEIN (FRAGMENT).
GN   CWT1.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=WHOLE EMBRYO;
RX   MEDLINE=96068905; PubMed=7478606;
RA   Kent J., Coriat A.M., Sharpe P.T., Hastie N., Van Heyningen V.;
RT   "The evolution of WT1 sequence and expression pattern in the
RT   vertebrates.";
RL   Oncogene 11:1781-1792(1995).
DR   EMBL; X85731; CAA59736.1; -.
DR   HSSP; P08046; 1A1G.
DR   InterPro; IPR000822; -.
DR   InterPro; IPR000976; -.
DR   Pfam; PF00096; zf-C2H2; 3.
DR   Pfam; PF02165; WT1; 2.
DR   PRINTS; PR00049; WILMSTUMOUR.
DR   PROSITE; PS00028; ZINC_FINGER_C2H2; 3.
DR   SMART; SM00355; ZnF_C2H2; 1.
KW   DNA-binding; Metal-binding; Zinc-finger.
FT   NON_TER     392    392
SQ   SEQUENCE   392 AA;  43869 MW;  761F01D350E4EDBF CRC64;


  Query Match             78.4%;  Score 40;  DB 13;  Length 392;
  Best Local Similarity   77.8%;  Pred. No. 9.5;
  Matches    7;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 RMFPNAPYL 9
           ||||  |||
Db     109 RMFPQRPYL 117

RESULT   6
Q9I8A1
ID   Q9I8A1      PRELIMINARY;      PRT;   414 AA.
AC   Q9I8A1;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   WT1(-KTS) PROTEIN.
GN   WT1(-KTS).
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=OVARY;
RA   Kudo T.;
RT   "Chicken counterpart of Wilms' tumor suppressor gene 1.";
RL   Submitted (OCT-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB033633; BAA94793.1; -.
DR   InterPro; IPR000822; -.
DR   InterPro; IPR000976; -.
DR   Pfam; PF00096; zf-C2H2; 4.
DR   Pfam; PF02165; WT1; 2.
DR   PRINTS; PR00049; WILMSTUMOUR.
DR   PROSITE; PS00028; ZINC_FINGER_C2H2; 4.
KW   DNA-binding; Metal-binding; Zinc-finger.
SQ   SEQUENCE   414 AA;  46641 MW;  17EB22F8B428A2EF CRC64;

  Query Match             78.4%;  Score 40;  DB 13;  Length 414;
  Best Local Similarity   77.8%;  Pred. No. 10;
  Matches    7;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 RMFPNAPYL 9
           ||||  |||
Db     109 RMFPQRPYL 117

RESULT   7
Q9I8A0
ID   Q9I8A0      PRELIMINARY;      PRT;   417 AA.
AC   Q9I8A0;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   WT1(+KTS) PROTEIN.
GN   WT1(+KTS).
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=OVARY;
RA   Kudo T.;
RT   "Chicken counterpart of Wilms' tumor suppressor gene 1.";
RL   Submitted (OCT-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB033634; BAA94794.1; -.
DR   InterPro; IPR000822; -.
DR   InterPro; IPR000976; -.
DR   Pfam; PF00096; zf-C2H2; 4.
DR   Pfam; PF02165; WT1; 2.
DR   PRINTS; PR00049; WILMSTUMOUR.
DR   PROSITE; PS00028; ZINC_FINGER_C2H2; 4.
KW   DNA-binding; Metal-binding; Zinc-finger.
SQ   SEQUENCE   417 AA;  46957 MW;  47BBF7F6448E7F6C CRC64;


  Query Match             78.4%;  Score 40;  DB 13;  Length 417;
  Best Local Similarity   77.8%;  Pred. No. 10;
  Matches    7;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 RMFPNAPYL 9
           ||||  |||
Db     109 RMFPQRPYL 117

RESULT   8
Q9PUT7
ID   Q9PUT7      PRELIMINARY;      PRT;   419 AA.
AC   Q9PUT7;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   WILMS' TUMOR SUPPRESSOR.
OS   Brachydanio rerio (Zebrafish) (Zebra danio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
OC   Cypriniformes; Cyprinidae; Rasborinae; Danio.
OX   NCBI_TaxID=7955;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Smith S.I., Down M., Power M., Boyd A.W.;
RT   "Isolation and characterization of a cDNA encoding zebrafish (Danio
RT   rerio) WT-1.";
RL   Submitted (APR-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF144550; AAF00123.1; -.
```

```
DR   HSSP; P08046; 1AAY.
DR   InterPro; IPR000822; -.
DR   InterPro; IPR000976; -.
DR   Pfam; PF00096; zf-C2H2; 4.
DR   Pfam; PF02165; WT1; 3.
DR   PRINTS; PR00049; WILMSTUMOUR.
DR   PROSITE; PS00028; ZINC_FINGER_C2H2; 4.
DR   SMART; SM00355; ZnF_C2H2; 1.
KW   DNA-binding; Metal-binding; Zinc-finger.
SQ   SEQUENCE   419 AA;  46925 MW;  606ADFEDA619EECD CRC64;

  Query Match             78.4%;  Score 40;  DB 13;  Length 419;
  Best Local Similarity   77.8%;  Pred. No. 10;
  Matches    7;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 RMFPNAPYL 9
          ||| |:|||
Db    111 RMFSNSPYL 119


RESULT   9
O83076
ID   O83076      PRELIMINARY;      PRT;   203 AA.
AC   O83076;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   CONSERVED HYPOTHETICAL INTEGRAL MEMBRANE PROTEIN.
GN   TP0033.
OS   Treponema pallidum.
OC   Bacteria; Spirochaetales; Spirochaetaceae; Treponema.
OX   NCBI_TaxID=160;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NICHOLS;
RX   MEDLINE=98332770; PubMed=9665876;
RA   Fraser C.M., Norris S.J., Weinstock G.M., White O., Sutton G.G.,
RA   Dodson R., Gwinn M., Hickey E.K., Clayton R., Ketchum K.A.,
RA   Sodergren E., Hardham J.M., McLeod M.P., Salzberg S., Peterson J.,
RA   Khalak H., Richardson D., Howell J.K., Chidambaram M., Utterback T.,
RA   McDonald L., Artiach P., Bowman C., Cotton M.D., Fujii C., Garland S.,
RA   Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O.,
RA   Venter J.C.;
RT   "Complete genome sequence of Treponema pallidum, the syphilis
RT   spirochete.";
RL   Science 281:375-388(1998).
DR   EMBL; AE001188; AAC65028.1; -.
DR   TIGR; TP0033; -.
SQ   SEQUENCE   203 AA;  23998 MW;  E4544E3457A063EA CRC64;

  Query Match             76.5%;  Score 39;  DB 2;  Length 203;
  Best Local Similarity   85.7%;  Pred. No. 7.4;
  Matches    6;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      3 FPNAPYL 9
          ||||||:
Db     63 FPNAPYI 69


RESULT  10
Q9IBF0
ID   Q9IBF0      PRELIMINARY;      PRT;   392 AA.
AC   Q9IBF0;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   WILMS' TUMOR PROTEIN.
GN   EWT1.
OS   Anguilla japonica (Japanese eel).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Anguilliformes; Anguilloidei;
OC   Anguillidae; Anguilla.
OX   NCBI_TaxID=7937;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=KIDNEY;
RX   MEDLINE=20183686; PubMed=10717475;
RA   Nakatsuru Y., Minami K., Yoshikawa A., Zhu J.J., Oda H., Masahito P.,
RA   Okamoto N., Nakamura Y., Ishikawa T.;
RT   "Eel WT1 sequence and expression in spontaneous nephroblastomas in
RT   Japanese eel.";
RL   Gene 245:245-251(2000).
DR   EMBL; AB030741; BAA90558.1; -.
DR   InterPro; IPR000822; -.
DR   InterPro; IPR000976; -.
DR   Pfam; PF00096; zf-C2H2; 4.
DR   Pfam; PF02165; WT1; 2.
DR   PRINTS; PR00049; WILMSTUMOUR.
DR   PROSITE; PS00028; ZINC_FINGER_C2H2; 4.
KW   DNA-binding; Metal-binding; Zinc-finger.
SQ   SEQUENCE   392 AA;  43892 MW;  DAE1A84828F43DF6 CRC64;

  Query Match             76.5%;  Score 39;  DB 13;  Length 392;
  Best Local Similarity   77.8%;  Pred. No. 15;
  Matches    7;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 RMFPNAPYL 9
          ||| | |||
Db    112 RMFANGPYL 120


RESULT  11
Q9W611
ID   Q9W611      PRELIMINARY;      PRT;   426 AA.
AC   Q9W611;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   WT1.
OS   Cynops pyrrhogaster (Japanese common newt).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Caudata; Salamandroidea; Salamandridae; Cynops.
OX   NCBI_TaxID=8330;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=TESTIS;
RA   Nakayama Y., Yamamoto T., Matsuda Y., Abe S.I.;
RT   "Cloning of cDNA for newt WT1 and the differential expression during
RT   spermatogenesis of the Japanese newt, Cynops pyrrhogaster.";
RL   Submitted (MAY-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB013888; BAA76399.1; -.
DR   HSSP; P08046; 1AAY.
DR   InterPro; IPR000822; -.
DR   InterPro; IPR000976; -.
DR   Pfam; PF00096; zf-C2H2; 4.
DR   Pfam; PF02165; WT1; 2.
DR   PRINTS; PR00049; WILMSTUMOUR.
DR   PROSITE; PS00028; ZINC_FINGER_C2H2; 4.
DR   SMART; SM00355; ZnF_C2H2; 1.
KW   DNA-binding; Metal-binding; Zinc-finger.
SQ   SEQUENCE   426 AA;  47590 MW;  46A0B158A9F79C6C CRC64;

  Query Match             76.5%;  Score 39;  DB 13;  Length 426;
  Best Local Similarity   77.8%;  Pred. No. 16;
  Matches    7;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 RMFPNAPYL 9
          ||| | |||
Db    116 RMFSNGPYL 124


RESULT  12
```

```
Q9I4D6
ID   Q9I4D6      PRELIMINARY;      PRT;   205 AA.
AC   Q9I4D6;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   PROBABLE HYDROLASE.
GN   PA1202.
OS   Pseudomonas aeruginosa.
OC   Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC   Pseudomonas.
OX   NCBI_TaxID=287;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PAO1;
RX   MEDLINE=20437337; PubMed=10984043;
RA   Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warrener P.,
RA   Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA   Garber R.L., Goltry L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
RA   Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA   Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA   Reizer J., Saier M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT   "Complete genome sequence of Pseudomonas aeruginosa PA01, an
RT   opportunistic pathogen.";
RL   Nature 406:959-964(2000).
DR   EMBL; AE004550; AAG04591.1; -.
DR   InterPro; IPR000868; -.
DR   Pfam; PF00857; Isochorismatase; 1.
KW   Hydrolase.
SQ   SEQUENCE   205 AA;  22471 MW;  B160D2F9BA4B625F CRC64;

  Query Match             74.5%;  Score 38;  DB 2;  Length 205;
  Best Local Similarity   75.0%;  Pred. No. 12;
  Matches    6;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        2 MFPNAPYL 9
            ||| |||:
Db       75 MFPQAPYI 82


RESULT  13
Q9FZ97
ID   Q9FZ97      PRELIMINARY;      PRT;   581 AA.
AC   Q9FZ97;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   F3H9.11 PROTEIN.
GN   F3H9.11.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC   Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Federspiel N.A., Palm C.J., Conway A.B., Conn L., Hansen N.F.,
RA   Altafi H., Nguyen M., Lam B., Southwick A., Miranda M., Brooks S.,
RA   Buehler E., Chao Q., Chin C., Chiou J., Choi E., Gonzalez A.,
RA   Howng B., Johnson-Hopson C., Khan S., Kim C., Koo T., Lee J.M.,
RA   Lenz C., Liu A., Liu S., Mukharsky N., Pham P., Sakano H., Shinn P.,
RA   Toriumi M., Vaysberg M., Yu G., Ecker J., Theologis A., Davis R.W.;
RL   Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AC021044; AAF98431.1; -.
SQ   SEQUENCE   581 AA;  65578 MW;  78E72252332F91AC CRC64;

  Query Match             74.5%;  Score 38;  DB 10;  Length 581;
  Best Local Similarity   87.5%;  Pred. No. 34;
  Matches    7;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        1 RMFPNAPY 8
            |||||| |
Db      349 RMFPNARY 356


RESULT  14
Q9N6S5
ID   Q9N6S5      PRELIMINARY;      PRT;   662 AA.
AC   Q9N6S5;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   511.2.
GN   511.2.
OS   Leishmania major.
OC   Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX   NCBI_TaxID=5664;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=FRIEDLIN;
RA   Myler P.J., Sisk E., Hixson G., Kiser P., Rickel E., Hassebrock M.,
RA   Cawthra J., Marsolini F., Sunkin S., Stuart K.D., Cunningham M.,
RA   Beverley S.;
RL   Submitted (JUN-2000) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=FRIEDLIN;
RA   Myler P.J.;
RL   Submitted (MAY-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AC005801; AAF77196.1; -.
DR   EMBL; AC009603; AAF70534.1; -.
SQ   SEQUENCE   662 AA;  68623 MW;  0B8C1D89A93891DB CRC64;

  Query Match             74.5%;  Score 38;  DB 5;  Length 662;
  Best Local Similarity   66.7%;  Pred. No. 39;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        1 RMFPNAPYL 9
            | |||||::
Db       33 RRFPNAPHM 41


RESULT  15
Q9W198
ID   Q9W198      PRELIMINARY;      PRT;  2175 AA.
AC   Q9W198;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   CG3363 PROTEIN.
GN   CG3363.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BERKELEY;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
```

```
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
DR   EMBL; AE003463; AAF47179.1; -.
DR   FlyBase; FBgn0034987; CG3363.
SQ   SEQUENCE   2175 AA;  241962 MW;  9722F6832021D758 CRC64;

  Query Match            74.5%;  Score 38;  DB 5;  Length 2175;
  Best Local Similarity   75.0%;  Pred. No. 1.3e+02;
  Matches    6;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 RMFPNAPY 8
            : ||||||
Db   1623 QFFPNAPY 1630


Search completed: October 23, 2001, 13:27:54
Job time: 407 sec
```

```
                       GenCore version 4.5
             Copyright (c) 1993 - 2000  Compugen Ltd.

OM protein - protein search, using sw model

Run on:         October 23, 2001, 13:21:37 ; Search time 17.76 Seconds
                                             (without alignments)
                                             17.359 Million cell updates/sec

Title:          US-09-625-963-1
Perfect score:  51
Sequence:       1 RMFPNAPYL 9

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched:       93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters:      93435

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       SwissProt_39:*

     Pred. No. is the number of results predicted by chance to have a
     score greater than or equal to the score of the result being printed,
     and is derived by analysis of the total score distribution.

                              SUMMARIES
                %
Result        Query
  No.   Score Match Length DB  ID                  Description
-------------------------------------------------------------------------
    1     51  100.0    448  1  WT1_RAT             P49952 rattus norv
    2     51  100.0    449  1  WT1_HUMAN           P19544 homo sapien
    3     51  100.0    449  1  WT1_MOUSE           P22561 mus musculu
    4     51  100.0    449  1  WT1_PIG             O62651 sus scrofa
    5     40   78.4    262  1  YFIO_HAEIN          P44553 haemophilus
    6     37   72.5    468  1  IFT2_CRILO          Q60462 cricetulus
    7     37   72.5    854  1  UN33_CAEEL          Q01630 caenorhabdi
    8     37   72.5   1403  1  BIR1_HUMAN          Q13075 homo sapien
    9     36   70.6    217  1  GTM1_HUMAN          P09488 homo sapien
   10     36   70.6    217  1  GTM1_MOUSE          P10649 mus musculu
   11     36   70.6    217  1  GTM1_RAT            P04905 rattus norv
   12     36   70.6    217  1  GTM2_HUMAN          P28161 homo sapien
   13     36   70.6    217  1  GTM2_MOUSE          P15626 mus musculu
   14     36   70.6    217  1  GTM2_RAT            P08010 rattus norv
   15     36   70.6    217  1  GTM3_MOUSE          P19639 mus musculu
   16     36   70.6    217  1  GTM3_RAT            P08009 rattus norv
   17     36   70.6    217  1  GTM5_HUMAN          P46439 homo sapien
   18     36   70.6    217  1  GTMU_CAVPO          P16413 cavia porce
   19     36   70.6    217  1  GTMU_CRILO          Q00285 cricetulus
   20     36   70.6    217  1  GTMU_MESAU          P30116 mesocricetu
   21     36   70.6    217  1  GTMU_RABIT          P46409 oryctolagus
   22     36   70.6    218  1  GTM4_HUMAN          Q03013 homo sapien
   23     36   70.6    219  1  GTM2_CHICK          P20136 gallus gall
   24     36   70.6    224  1  GTM5_MOUSE          P48774 mus musculu
   25     36   70.6    225  1  GTM3_HUMAN          P21266 homo sapien
   26     36   70.6    226  1  BSH_DROME           Q04787 drosophila
   27     36   70.6    358  1  PIAP_PIG            O62640 sus scrofa
   28     36   70.6    370  1  TPS1_HUMAN          O60507 homo sapien
   29     36   70.6    370  1  TPS1_MOUSE          O70281 mus musculu
   30     36   70.6    381  1  RSA1_YEAST          Q08932 saccharomyc
   31     36   70.6    382  1  KSE1_ECOLI          P42501 escherichia
   32     36   70.6    382  1  KSE5_ECOLI          P42214 escherichia
   33     36   70.6    487  1  C8B1_ESCCA          O64899 eschscholzi
   34     36   70.6    488  1  C8B2_ESCCA          O64900 eschscholzi
   35     36   70.6    509  1  C4A1_DROME          Q9vms9 drosophila
   36     36   70.6    618  1  BIR3_HUMAN          Q13490 homo sapien
   37     35   68.6    273  1  TRY6_ANOGA          P35040 anopheles g
   38     35   68.6    380  1  TPSA_CAEEL          O77081 caenorhabdi
   39     35   68.6    500  1  C912_ARATH          O65790 arabidopsis
   40     34   66.7    902  1  ST14_MOUSE          P56677 mus musculu
   41     34   66.7   1097  1  CCT_DROME           O96433 drosophila
   42     33   64.7    110  1  YLXM_STRMU          P96468 streptococc
   43     33   64.7    138  1  RL16_BORBU          O51438 borrelia bu
   44     33   64.7    139  1  RL16_TREPA          O83226 treponema p
   45     33   64.7    142  1  RL16_AQUAE          O66438 aquifex aeo

                              ALIGNMENTS

RESULT   1
WT1_RAT
ID   WT1_RAT        STANDARD;      PRT;   448 AA.
AC   P49952;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   WILMS' TUMOR PROTEIN HOMOLOG.
GN   WT1 OR WT-1.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SPRAGUE-DAWLEY; TISSUE=Kidney;
RX   MEDLINE=93046155; PubMed=1330293;
RA   Sharma P.M., Yang X., Bowman M., Roberts V., Sukumar S.;
RT   "Molecular cloning of rat Wilms' tumor complementary DNA and a study
RT   of messenger RNA expression in the urogenital system and the brain.";
RL   Cancer Res. 52:6407-6412(1992).
CC   -!- FUNCTION: POTENTIAL ROLE IN TRANSCRIPTIONAL REGULATION. RECOGNIZES
CC       AND BINDS TO THE DNA SEQUENCE 5'-CGCCCCCGC-3'.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR.
CC   -!- ALTERNATIVE PRODUCTS: TWO ALTERNATIVE SPLICE SITES EXISTS.
CC   -!- TISSUE SPECIFICITY: KIDNEY (BY SIMILARITY).
CC   -!- DEVELOPMENTAL STAGE: EXPRESSED DURING KIDNEY DEVELOPMENT.
CC   -!- SIMILARITY: BELONGS TO THE EGR FAMILY OF C2H2-TYPE ZINC-FINGER
CC       PROTEINS.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; X69716; CAA49373.1; -.
DR   TRANSFAC; T02352; -.
DR   InterPro; IPR000822; -.
DR   InterPro; IPR000976; -.
DR   Pfam; PF00096; zf-C2H2; 4.
DR   PRINTS; PR00048; ZINCFINGER.
DR   PRINTS; PR00049; WILMSTUMOUR.
DR   PROSITE; PS00028; ZINC_FINGER_C2H2_1; 4.
DR   PROSITE; PS50157; ZINC_FINGER_C2H2_2; 4.
KW   Zinc-finger; Metal-binding; DNA-binding; Repeat; Nuclear protein;
KW   Transcription regulation; Alternative splicing; Anti-oncogene.
FT   DOMAIN       27     82       PRO-RICH.
FT   DOMAIN      322    437       ZINC FINGERS.
FT   ZN_FING     322    346       C2H2-TYPE.
FT   ZN_FING     352    376       C2H2-TYPE.
FT   ZN_FING     382    404       C2H2-TYPE.
FT   ZN_FING     413    437       C2H2-TYPE.
FT   VARSPLIC    249    265       MISSING (IN ISOFORM 2 AND ISOFORM 3).
```

```
FT   VARSPLIC     407    409       MISSING (IN ISOFORM 2 AND ISOFORM 4).
SQ   SEQUENCE   448 AA;  49193 MW;  329AC9AC1FF73F76 CRC64;


  Query Match             100.0%;  Score 51;  DB 1;  Length 448;
  Best Local Similarity   100.0%;  Pred. No. 0.02;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 RMFPNAPYL 9
            |||||||||
Db      125 RMFPNAPYL 133


RESULT   2
WT1_HUMAN
ID   WT1_HUMAN      STANDARD;      PRT;   449 AA.
AC   P19544;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-AUG-1991 (Rel. 19, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   WILMS' TUMOR PROTEIN (WT33).
GN   WT1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Fetal kidney;
RX   MEDLINE=90158822; PubMed=2154702;
RA   Gessler M., Poustka A., Cavenee W., Neve R.L., Orkin S.H.,
RA   Bruns G.A.P.;
RT   "Homozygous deletion in Wilms tumours of a zinc-finger gene
RT   identified by chromosome jumping.";
RL   Nature 343:774-778(1990).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Placenta;
RX   MEDLINE=92052142; PubMed=1658787;
RA   Haber D.A., Sohn R.L., Buckler A.J., Pelletier J., Call K.M.,
RA   Housman D.E.;
RT   "Alternative splicing and genomic structure of the Wilms tumor gene
RT   WT1.";
RL   Proc. Natl. Acad. Sci. U.S.A. 88:9618-9622(1991).
RN   [3]
RP   SEQUENCE OF 85-449 FROM N.A.
RX   MEDLINE=90150277; PubMed=2154335;
RA   Call K.M., Glaser T., Ito C.Y., Buckler A.J., Pelletier J.,
RA   Haber D.A., Rose E.A., Kral A., Yeger H., Lewis W.H., Jones C.,
RA   Housman D.E.;
RT   "Isolation and characterization of a zinc finger polypeptide gene at
RT   the human chromosome 11 Wilms' tumor locus.";
RL   Cell 60:509-520(1990).
RN   [4]
RP   IDENTIFICATION OF START CODON AND ALTERNATIVE SPLICE SITES.
RX   MEDLINE=91141522; PubMed=1671709;
RA   Buckler A.J., Pelletier J., Haber D.A., Glaser T., Housman D.E.;
RT   "Isolation, characterization, and expression of the murine Wilms'
RT   tumor gene (WT1) during kidney development.";
RL   Mol. Cell. Biol. 11:1707-1712(1991).
RN   [5]
RP   VARIANT WT CYS-366.
RX   MEDLINE=92279213; PubMed=1317572;
RA   Little M.H., Prosser J., Condie A., Smith P.J., van Heyningen V.,
RA   Hastie N.D.;
RT   "Zinc finger point mutations within the WT1 gene in Wilms tumor
RT   patients.";
RL   Proc. Natl. Acad. Sci. U.S.A. 89:4791-4795(1992).
RN   [6]
RP   VARIANTS DDS.
RX   MEDLINE=92005721; PubMed=1655284;
RA   Pelletier J., Bruening W., Kashtan C.E., Mauer S.M., Manivel J.C.,
RA   Striegel J.E., Houghton D.C., Junien C., Habib R., Fouser L.,
RA   Fine R.N., Silverman B.L., Haber D.A., Housman D.;
RT   "Germline mutations in the Wilms' tumor suppressor gene are
RT   associated with abnormal urogenital development in Denys-Drash
RT   syndrome.";
RL   Cell 67:437-447(1991).
RN   [7]
RP   VARIANTS DDS.
RX   MEDLINE=93265053; PubMed=1338906;
RA   Baird P.N., Santos A., Groves N., Jadresic L., Cowell J.K.;
RT   "Constitutional mutations in the WT1 gene in patients with
RT   Denys-Drash syndrome.";
RL   Hum. Mol. Genet. 1:301-305(1992).
RN   [8]
RP   VARIANTS DDS.
RX   MEDLINE=93271983; PubMed=8388765;
RA   Little M.H., Williamson K.A., Mannens M., Kelsey A., Gosden C.,
RA   Hastie N., van Heyningen V.;
RT   "Evidence that WT1 mutations in Denys-Drash syndrome patients may act
RT   in a dominant-negative fashion.";
RL   Hum. Mol. Genet. 2:259-264(1993).
RN   [9]
RP   VARIANT MESOTHELIOMA GLY-273.
RX   MEDLINE=94004972; PubMed=8401592;
RA   Park S., Schalling M., Bernard A., Maheswaran S., Shipley G.C.,
RA   Roberts D., Fletcher J., Shipman R., Rheinwald J., Demetri G.,
RA   Griffin J., Minden M., Housman D.E., Haber D.A.;
RT   "The Wilms tumour gene WT1 is expressed in murine mesoderm-derived
RT   tissues and mutated in a human mesothelioma.";
RL   Nat. Genet. 4:415-420(1993).
RN   [10]
RP   VARIANTS WT SER-181 AND ALA-253.
RX   MEDLINE=97268681; PubMed=9108089;
RA   Schumacher V., Schneider S., Figge A., Wildhardt G., Harms D.,
RA   Schmidt D., Weirich A., Ludwig R., Royer-Pokora B.;
RT   "Correlation of germ-line mutations and two-hit inactivation of the
RT   WT1 gene with Wilms tumors of stromal-predominant histology.";
RL   Proc. Natl. Acad. Sci. U.S.A. 94:3972-3977(1997).
RN   [11]
RP   VARIANTS DMS TYR-377; LEU-383 AND ASN-396.
RX   MEDLINE=98198341; PubMed=9529364;
RA   Jeanpierre C., Denamur E., Henry I., Cabanis M.-O., Luce S.,
RA   Cecille A., Elion J., Peuchmaur M., Loirat C., Niaudet P.,
RA   Gubler M.-C., Junien C.;
RT   "Identification of constitutional WT1 mutations, in patients with
RT   isolated diffuse mesangial sclerosis, and analysis of
RT   genotype/phenotype correlations by use of a computerized mutation
RT   database.";
RL   Am. J. Hum. Genet. 62:824-833(1998).
RN   [12]
RP   REVIEW.
RX   MEDLINE=92207913; PubMed=1313285;
RA   Haber D.A., Buckler A.J.;
RT   "WT1: a novel tumor suppressor gene inactivated in Wilms' tumor.";
RL   New Biol. 4:97-106(1992).
RN   [13]
RP   REVIEW.
RX   MEDLINE=93345769; PubMed=8393820;
RA   Rauscher F.J. III;
RT   "The WT1 Wilms tumor gene product: a developmentally regulated
RT   transcription factor in the kidney that functions as a tumor
RT   suppressor.";
RL   FASEB J. 7:896-903(1993).
CC   -!- FUNCTION: POTENTIAL ROLE IN TRANSCRIPTIONAL REGULATION. RECOGNIZES
CC       AND BINDS TO THE DNA SEQUENCE 5'-CGCCCCCGC-3'.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR.
CC   -!- ALTERNATIVE PRODUCTS: TWO ALTERNATIVE SPLICE SITES EXISTS.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN THE KIDNEY AND A SUBSET OF
CC       HEMATOPOIETIC CELLS.
CC   -!- DISEASE: WILM'S TUMOR (WT) IS AN EMBRYONAL MALIGNANCY OF THE
CC       KIDNEY THAT AFFECTS APPROXIMATELY 1 IN 10,000 INFANTS AND YOUNG
CC       CHILDREN. IT OCCURS BOTH IN SPORADIC AND HEREDITARY FORMS.
CC       INACTIVATION OF WT1 IS ONE OF THE CAUSES OF WILM'S TUMOR.
CC   -!- DISEASE: PATIENTS WITH DENYS-DRASH SYNDROME (DDS), WHICH IS
```

```
CC       CHARACTERIZED BY A TYPICAL NEPHROPATHY AND GENITAL ABNORMALITIES,
CC       HAVE DEFECTS IN THE WT1 GENE.
CC   -!- DISEASE: DEFECTS IN WT1 ARE ALSO A CAUSE OF DIFFUSE MESANGIAL
CC       SCLEROSIS (DMS), A FORM A DDS.
CC   -!- SIMILARITY: BELONGS TO THE EGR FAMILY OF C2H2-TYPE ZINC-FINGER
CC       PROTEINS.
CC   -!- DATABASE: NAME=WT1; NOTE=WT1 mutation database;
CC       WWW="http://www.umd.necker.fr:2003/".
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; X51630; CAA35956.1; ALT_INIT.
DR   EMBL; M80232; AAA61299.1; -.
DR   EMBL; M80217; AAA61299.1; JOINED.
DR   EMBL; M80218; AAA61299.1; JOINED.
DR   EMBL; M80219; AAA61299.1; JOINED.
DR   EMBL; M80220; AAA61299.1; JOINED.
DR   EMBL; M80221; AAA61299.1; JOINED.
DR   EMBL; M80228; AAA61299.1; JOINED.
DR   EMBL; M80229; AAA61299.1; JOINED.
DR   EMBL; M80231; AAA61299.1; JOINED.
DR   EMBL; M30393; AAA36810.1; -.
DR   PIR; A34673; A34673.
DR   PIR; S08273; S08273.
DR   TRANSFAC; T00899; -.
DR   MIM; 194070; -.
DR   MIM; 194080; -.
DR   MIM; 256370; -.
DR   InterPro; IPR000822; -.
DR   InterPro; IPR000976; -.
DR   Pfam; PF00096; zf-C2H2; 4.
DR   PRINTS; PR00048; ZINCFINGER.
DR   PRINTS; PR00049; WILMSTUMOUR.
DR   PROSITE; PS00028; ZINC_FINGER_C2H2_1; 4.
DR   PROSITE; PS50157; ZINC_FINGER_C2H2_2; 4.
KW   Zinc-finger; Metal-binding; DNA-binding; Repeat; Nuclear protein;
KW   Transcription regulation; Alternative splicing; Anti-oncogene;
KW   Disease mutation.
FT   DOMAIN       27     83       PRO-RICH.
FT   DOMAIN      323    438       ZINC FINGERS.
FT   ZN_FING     323    347       C2H2-TYPE.
FT   ZN_FING     353    377       C2H2-TYPE.
FT   ZN_FING     383    405       C2H2-TYPE.
FT   ZN_FING     414    438       C2H2-TYPE.
FT   VARSPLIC    250    266       MISSING (IN ISOFORM 2 AND ISOFORM 3).
FT   VARSPLIC    408    410       MISSING (IN ISOFORM 2 AND ISOFORM 4).
FT   VARIANT     181    181       P -> S (IN WT).
FT                                /FTId=VAR_007739.
FT   VARIANT     223    223       S -> N (IN WT).
FT                                /FTId=VAR_007740.
FT   VARIANT     253    253       G -> A (IN WT).
FT                                /FTId=VAR_007741.
FT   VARIANT     273    273       S -> G (IN MESOTHELIOMA).
FT                                /FTId=VAR_007742.
FT   VARIANT     330    330       C -> Y (IN DDS).
FT                                /FTId=VAR_007743.
FT   VARIANT     360    360       C -> G (IN DDS).
FT                                /FTId=VAR_007744.
FT   VARIANT     366    366       R -> C (IN WT).
FT                                /FTId=VAR_007745.
FT   VARIANT     366    366       R -> H (IN DDS).
FT                                /FTId=VAR_007746.
FT   VARIANT     373    373       H -> Q (IN DDS).

  Query Match              100.0%;  Score 51;  DB 1;  Length 449;
  Best Local Similarity   100.0%;  Pred. No. 0.02;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RMFPNAPYL 9
              |||||||||
Db        126 RMFPNAPYL 134

RESULT 3
WT1_MOUSE
ID   WT1_MOUSE      STANDARD;      PRT;   449 AA.
AC   P22561;
DT   01-AUG-1991 (Rel. 19, Created)
DT   01-AUG-1991 (Rel. 19, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   WILMS' TUMOR PROTEIN HOMOLOG.
GN   WT1 OR WT-1.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91141522; PubMed=1671709;
RA   Buckler A.J., Pelletier J., Haber D.A., Glaser T., Housman D.E.;
RT   "Isolation, characterization, and expression of the murine Wilms'
RT   tumor gene (WT1) during kidney development.";
RL   Mol. Cell. Biol. 11:1707-1712(1991).
CC   -!- FUNCTION: POTENTIAL ROLE IN TRANSCRIPTIONAL REGULATION. RECOGNIZES
CC       AND BINDS TO THE DNA SEQUENCE 5'-CGCCCCCGC-3'.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR.
CC   -!- ALTERNATIVE PRODUCTS: TWO ALTERNATIVE SPLICE SITES EXISTS.
CC   -!- TISSUE SPECIFICITY: KIDNEY.
CC   -!- DEVELOPMENTAL STAGE: EXPRESSED DURING KIDNEY DEVELOPMENT.
CC   -!- SIMILARITY: BELONGS TO THE EGR FAMILY OF C2H2-TYPE ZINC-FINGER
CC       PROTEINS.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; M55512; AAA40573.1; -.
DR   PIR; A39692; A39692.
DR   TRANSFAC; T02351; -.
DR   MGD; MGI:98968; Wt1.
DR   InterPro; IPR000822; -.
DR   InterPro; IPR000976; -.
DR   Pfam; PF00096; zf-C2H2; 4.
DR   PRINTS; PR00048; ZINCFINGER.
DR   PRINTS; PR00049; WILMSTUMOUR.
DR   PROSITE; PS00028; ZINC_FINGER_C2H2_1; 4.
DR   PROSITE; PS50157; ZINC_FINGER_C2H2_2; 4.
KW   Zinc-finger; Metal-binding; DNA-binding; Repeat; Nuclear protein;
KW   Transcription regulation; Alternative splicing; Anti-oncogene.
FT   DOMAIN       28     83       PRO-RICH.
FT   DOMAIN      323    438       ZINC FINGERS.
FT   ZN_FING     323    347       C2H2-TYPE.
FT   ZN_FING     353    377       C2H2-TYPE.
FT   ZN_FING     383    405       C2H2-TYPE.
FT   ZN_FING     414    438       C2H2-TYPE.
FT   VARSPLIC    250    266       MISSING (IN ISOFORM 2 AND ISOFORM 3).
FT   VARSPLIC    408    410       MISSING (IN ISOFORM 2 AND ISOFORM 4).
SQ   SEQUENCE   449 AA;  49246 MW;  962381E9C8D7A380 CRC64;

  Query Match              100.0%;  Score 51;  DB 1;  Length 449;
  Best Local Similarity   100.0%;  Pred. No. 0.02;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RMFPNAPYL 9
              |||||||||
```

```
Db    126 RMFPNAPYL 134


RESULT   4
WT1_PIG
ID   WT1_PIG        STANDARD;      PRT;   449 AA.
AC   O62651;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   WILMS' TUMOR PROTEIN HOMOLOG.
GN   WT1.
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX   NCBI_TaxID=9823;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=LWD; TISSUE=Kidney;
RX   MEDLINE=98267201; PubMed=9602131;
RA   Tsurutani N., Oda H., Nakatsuru Y., Imai Y., Zhang S., Ueno Y.,
RA   Ishikawa T.;
RT   "cDNA cloning and developmental expression of the porcine homologue
RT   of WT1.";
RL   Gene 211:215-220(1998).
CC   -!- FUNCTION: POTENTIAL ROLE IN TRANSCRIPTIONAL REGULATION. RECOGNIZES
CC       AND BINDS TO THE DNA SEQUENCE 5'-CGCCCCCGC-3'.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR.
CC   -!- ALTERNATIVE PRODUCTS: TWO ALTERNATIVE SPLICE SITES EXISTS.
CC   -!- DEVELOPMENTAL STAGE: EXPRESSED DURING KIDNEY DEVELOPMENT.
CC   -!- SIMILARITY: BELONGS TO THE EGR FAMILY OF C2H2-TYPE ZINC-FINGER
CC       PROTEINS.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AB010969; BAA28147.1; -.
DR   InterPro; IPR000822; -.
DR   InterPro; IPR000976; -.
DR   Pfam; PF00096; zf-C2H2; 4.
DR   PRINTS; PR00048; ZINCFINGER.
DR   PRINTS; PR00049; WILMSTUMOUR.
DR   PROSITE; PS00028; ZINC_FINGER_C2H2_1; 4.
DR   PROSITE; PS50157; ZINC_FINGER_C2H2_2; 4.
KW   Zinc-finger; Metal-binding; DNA-binding; Repeat; Nuclear protein;
KW   Transcription regulation; Alternative splicing; Anti-oncogene.
FT   DOMAIN       28     83       PRO-RICH.
FT   DOMAIN      323    438       ZINC FINGERS.
FT   ZN_FING     323    347       C2H2-TYPE.
FT   ZN_FING     353    377       C2H2-TYPE.
FT   ZN_FING     383    405       C2H2-TYPE.
FT   ZN_FING     414    438       C2H2-TYPE.
FT   VARSPLIC    249    265       MISSING (IN ISOFORM 2 AND ISOFORM 3) (BY
FT                                SIMILARITY).
FT   VARSPLIC    407    409       MISSING (IN ISOFORM 2 AND ISOFORM 4) (BY
FT                                SIMILARITY).
SQ   SEQUENCE   449 AA;  49166 MW;  9C3E557B96F5A7B3 CRC64;


  Query Match              100.0%;  Score 51;  DB 1;  Length 449;
  Best Local Similarity   100.0%;  Pred. No. 0.02;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 RMFPNAPYL 9
          |||||||||
Db    126 RMFPNAPYL 134
```

```
RESULT   5
YFIO_HAEIN
ID   YFIO_HAEIN     STANDARD;      PRT;   262 AA.
AC   P44553;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   PUTATIVE LIPOPROTEIN HI0177 PRECURSOR.
GN   HI0177.
OS   Haemophilus influenzae.
OC   Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC   Haemophilus.
OX   NCBI_TaxID=727;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=RD / KW20 / ATCC 51907;
RX   MEDLINE=95350630; PubMed=7542800;
RA   Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA   Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA   McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA   Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA   Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA   Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA   Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
RA   Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA   Venter J.C.;
RT   "Whole-genome random sequencing and assembly of Haemophilus
RT   influenzae Rd.";
RL   Science 269:496-512(1995).
RN   [2]
RP   IDENTIFICATION BY MASS SPECTROMETRY.
RX   MEDLINE=20137488; PubMed=10675023;
RA   Langen H., Takacs B., Evers S., Berndt P., Lahm H.W., Wipf B.,
RA   Gray C., Fountoulakis M.;
RT   "Two-dimensional map of the proteome of Haemophilus influenzae.";
RL   Electrophoresis 21:411-429(2000).
CC   -!- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A LIPID
CC       ANCHOR (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE COML FAMILY OF LIPOPROTEINS. STRONG,
CC       TO E.COLI YFIO.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; U32703; AAC21847.1; -.
DR   TIGR; HI0177; -.
KW   Membrane; Lipoprotein; Signal.
FT   SIGNAL        1     18       POTENTIAL.
FT   CHAIN        19    262       PUTATIVE LIPOPROTEIN HI0177.
FT   LIPID        19     19       N-ACYL DIGLYCERIDE (POTENTIAL).
SQ   SEQUENCE   262 AA;  29347 MW;  1DD9FFFE568D2B7B CRC64;


  Query Match              78.4%;  Score 40;  DB 1;  Length 262;
  Best Local Similarity    75.0%;  Pred. No. 1.4;
  Matches    6;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 RMFPNAPY 8
          |:|||:||
Db    151 RVFPNSPY 158


RESULT   6
IFT2_CRILO
ID   IFT2_CRILO     STANDARD;      PRT;   468 AA.
AC   Q60462;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
```

```
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   INTERFERON-INDUCED PROTEIN WITH TETRATRICOPEPTIDE REPEATS 2 (IFIT-2)
DE   (INTERFERON-INDUCED 54 KDA PROTEIN) (IFI-54K) (CL-54 K).
GN   IFIT2 OR IFI54.
OS   Cricetulus longicaudatus (Long-tailed hamster) (Chinese hamster).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC   Cricetulus.
OX   NCBI_TaxID=10030;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94170787; PubMed=8125096;
RA   Bluyssen J.A.R., Vlietstra R.J., van der Made A., Trapman J.;
RT   "The interferon-stimulated gene 54 K promoter contains two adjacent
RT   functional interferon-stimulated response elements of different
RT   strength, which act synergistically for maximal interferon-alpha
RT   inducibility.";
RL   Eur. J. Biochem. 220:395-402(1994).
CC   -!- INDUCTION: BY INTERFERONS.
CC   -!- SIMILARITY: BELONGS TO THE IFIT FAMILY.
CC   -!- SIMILARITY: CONTAINS TPR REPEATS.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; X77259; CAA54477.1; -.
KW   Repeat; TPR repeat; Interferon induction.
FT   REPEAT       51     85       TPR 1.
FT   REPEAT       87    128       TPR 2.
FT   REPEAT      157    172       TPR 3B.
FT   REPEAT      175    209       TPR 4.
FT   REPEAT      210    242       TPR 5.
FT   REPEAT      243    276       TPR 6.
FT   REPEAT      277    287       TPR 7A.
FT   REPEAT      312    327       TPR 7B.
FT   REPEAT      328    361       TPR 8.
FT   REPEAT      364    399       TPR 9.
FT   REPEAT      423    457       TPR 10.
SQ   SEQUENCE   468 AA;  55045 MW;  9D92B878F81B1E27 CRC64;


  Query Match             72.5%;  Score 37;  DB 1;  Length 468;
  Best Local Similarity   66.7%;  Pred. No. 9.5;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        1 RMFPNAPYL 9
            || ||:||:
Db      203 RMSPNSPYV 211


RESULT   7
UN33_CAEEL
ID   UN33_CAEEL     STANDARD;      PRT;   854 AA.
AC   Q01630;
DT   01-JUL-1993 (Rel. 26, Created)
DT   01-JUL-1993 (Rel. 26, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   UNC-33 PROTEIN.
GN   UNC-33.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RX   MEDLINE=93106371; PubMed=1468626;
RA   Li W., Herman R.K., Shaw J.E.;
```

```
RT   "Analysis of the Caenorhabditis elegans axonal guidance and outgrowth
RT   gene unc-33.";
RL   Genetics 132:675-689(1992).
CC   -!- FUNCTION: INVOLVED IN AXONAL GUIDANCE AND OUTGROWTH.
CC   -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC   -!- ALTERNATIVE PRODUCTS: THREE PUTATIVE POLYPEPTIDES MAY ARISE BY
CC       ALTERNATIVE SPLICING.
CC   -!- DEVELOPMENTAL STAGE: UNC-33 PROTEINS ARE DISTRIBUTED EXCLUSIVELY
CC       WITHIN NEURONAL PROCESSES AFTER EARLY EMBRYOGENESIS.
CC   -!- SIMILARITY: BELONGS TO THE DEHYDROPYRIMIDINASE FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; Z14148; CAA78520.1; -.
DR   EMBL; Z14148; CAA78521.1; -.
DR   EMBL; Z14148; CAA78522.1; -.
DR   EMBL; Z14146; CAA78516.1; -.
DR   EMBL; Z14146; CAA78517.1; -.
DR   EMBL; Z14146; CAA78518.1; -.
DR   PIR; S24643; S24643.
DR   PIR; S24644; S24644.
DR   PIR; S33558; S33558.
DR   InterPro; IPR002195; -.
DR   Pfam; PF00744; Dihydroorotase; 1.
KW   Alternative splicing.
FT   VARSPLIC      1    175       MISSING (IN 72.1 KDA ISOFORM).
FT   VARSPLIC      1    331       MISSING (IN 55.5 KDA ISOFORM).
SQ   SEQUENCE   854 AA;  90819 MW;  A8073DDE251D2D77 CRC64;


  Query Match             72.5%;  Score 37;  DB 1;  Length 854;
  Best Local Similarity   66.7%;  Pred. No. 18;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        1 RMFPNAPYL 9
            |: ||:|||
Db      773 RLSPNSPYL 781


RESULT   8
BIR1_HUMAN
ID   BIR1_HUMAN     STANDARD;      PRT;  1403 AA.
AC   Q13075; Q13730; Q99796; O75857;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-OCT-2000 (Rel. 40, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   BACULOVIRAL IAP REPEAT-CONTAINING PROTEIN 1 (NEURONAL APOPTOSIS
DE   INHIBITORY PROTEIN).
GN   BIRC1 OR NAIP.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Fetal brain;
RX   MEDLINE=95112344; PubMed=7813013;
RA   Roy N., Mahadevan M.S., McLean M., Shutler G., Yaraghi Z.,
RA   Farahini R., Baird S., Besner-Johnston A., Lefebvre C., Kang X.,
RA   Salih M., Aubry H., Tamai K., Guan X., Ioannou P., Crawford T.O.,
RA   de Jong P.J., Surh L., Ikeda J., Korneluk R.G., Mackenzie A.;
RT   "The gene for neuronal apoptosis inhibitory protein is partially
RT   deleted in individuals with spinal muscular atrophy.";
RL   Cell 80:167-178(1995).
RN   [2]
RP   SEQUENCE FROM N.A., AND REVISIONS.
RC   TISSUE=Brain;
```

```
RX   MEDLINE=98163755; PubMed=9503025;
RA   Chen Q., Baird S.D., Mahadevan M., Besner-Johnston A., Farahani R.,
RA   Xuan J.-Y., Kang X., Lefebvre C., Ikeda J.-E., Korneluk R.G.,
RA   MacKenzie A.E.;
RT   "Sequence of a 131-kb region of 5q13.1 containing the spinal muscular
RT   atrophy candidate genes SMN and NAIP.";
RL   Genomics 48:121-127(1998).
RN   [3]
RP   SEQUENCE OF 386-623 FROM N.A.
RA   der Steege G., Draaijers T.G., Grootscholten P.M., Osinga J.,
RA   Anzevino R., Velona I., Brahe C., Scheffer.H., van Ommen G.J.B.,
RA   Buys C.H.C.M.;
RL   Submitted (MAY-1995) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE OF 222-1403 FROM N.A.
RA   Jones K., Graves T., McPherson J.;
RL   Submitted (JUN-1998) to the EMBL/GenBank/DDBJ databases.
RN   [5]
RP   FUNCTION.
RC   TISSUE=Liver;
RX   MEDLINE=96149249; PubMed=8552191;
RA   Liston P., Roy N., Tamai K., Lefebvre C., Baird S., Cherton-Horvat G.,
RA   Farahani R., McLean M., Ikeda J., Mackenzie A., Korneluk R.G.;
RT   "Suppression of apoptosis in mammalian cells by NAIP and a related
RT   family of IAP genes.";
RL   Nature 379:349-353(1996).
CC   -!- FUNCTION: PREVENTS MOTOR-NEURON APOPTOSIS INDUCED BY A VARIETY OF
CC       SIGNALS.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN MOTOR NEURONS, BUT NOT IN SENSORY
CC       NEURONS. FOUND IN LIVER AND PLACENTA, AND IN A LESSER EXTENT IN
CC       SPINAL CORD.
CC   -!- DISEASE: MUTATED OR DELETED FORMS OF NAIP HAVE BEEN FOUND IN
CC       INDIVIDUALS WITH SPINAL MUSCULAR ATROPHY TYPE I (SMA TYPE 1). SMAS
CC       ARE FATAL AUTOSOMAL RECESSIVE DISORDERS SUBCLASSIFIED AS TYPE 1
CC       (WERDNIG-HOFFMANN DISEASE), TYPE II (INTERMEDIATE FORM), AND TYPE
CC       III (WOHLFART-KUGELBERG-WELANDER DISEASE) BASED UPON THE AGE OF
CC       ONSET AND CLINICAL SEVERITY. THESE NEURODEGENERATIVE DISORDERS ARE
CC       CHARACTERIZED BY DEGENERATION OF LOWER MOTOR NEURONS, LEADING TO
CC       PROGRESSIVE PARALYSIS MUSCULAR ATROPHY. CONCERNS 1 IN 6000
CC       NEWBORNS.
CC   -!- SIMILARITY: CONTAINS 3 BIR REPEATS.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; U19251; AAC52045.1; -.
DR   EMBL; U80017; AAC52047.1; -.
DR   EMBL; U21913; AAA64504.1; -.
DR   EMBL; AC005031; AAC62261.1; -.
DR   MIM; 600355; -.
DR   InterPro; IPR001370; -.
DR   Pfam; PF00653; BIR; 3.
DR   PROSITE; PS01282; BIR_REPEAT_1; 3.
DR   PROSITE; PS50143; BIR_REPEAT_2; 3.
KW   Apoptosis; Repeat.
FT   REPEAT       60    127       BIR 1.
FT   REPEAT      159    227       BIR 2.
FT   REPEAT      278    345       BIR 3.
FT   CONFLICT    222    223       PK -> YR (IN REF. 4).
FT   CONFLICT    386    387       VP -> ST (IN REF. 3).
FT   CONFLICT    535    535       M -> V (IN REF. 3).
FT   CONFLICT    553    553       Y -> H (IN REF. 3).
FT   CONFLICT   1228   1231       MISSING (IN REF. 4).
SQ   SEQUENCE   1403 AA;  159613 MW;  566304C154DA5E64 CRC64;

  Query Match             72.5%;  Score 37;  DB 1;  Length 1403;
  Best Local Similarity   66.7%;  Pred. No. 31;
```

```
  Matches    6;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 RMFPNAPYL 9
           | ||| |:|
Db     337 RCFPNCPFL 345


RESULT   9
GTM1_HUMAN
ID   GTM1_HUMAN     STANDARD;      PRT;   217 AA.
AC   P09488;
DT   01-MAR-1989 (Rel. 10, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   GLUTATHIONE S-TRANSFERASE MU 1 (EC 2.5.1.18) (GSTM1-1) (HB SUBUNIT 4)
DE   (GTH4) (GSTM1A-1A) (GSTM1B-1B) (GST CLASS-MU).
GN   GSTM1 OR GST1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=88335606; PubMed=3419925;
RA   Dejong J.L., Chang C.M., Whang Peng J., Knutsen T., Tu C.-P.D.;
RT   "The human liver glutathione S-transferase gene superfamily:
RT   expression and chromosome mapping of an Hb subunit cDNA.";
RL   Nucleic Acids Res. 16:8541-8554(1988).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=89017184; PubMed=3174634;
RA   Seidegaard J., Vorachek W.R., Pero R.W., Pearson W.R.;
RT   "Hereditary differences in the expression of the human glutathione
RT   transferase active on trans-stilbene oxide are due to a gene
RT   deletion.";
RL   Proc. Natl. Acad. Sci. U.S.A. 85:7293-7297(1988).
RN   [3]
RP   SEQUENCE OF 1-188 FROM N.A.
RX   MEDLINE=93228631; PubMed=8471052;
RA   Zhong S., Spurr N.K., Hayes J.D., Wolf C.R.;
RT   "Deduced amino acid sequence, gene structure and chromosomal location
RT   of a novel human class Mu glutathione S-transferase, GSTM4.";
RL   Biochem. J. 291:41-50(1993).
RN   [4]
RP   SEQUENCE OF 59-117 FROM N.A.
RX   MEDLINE=90301515; PubMed=2362832;
RA   Comstock K.E., Sanderson B.J.S., Claflin G., Henner W.D.;
RT   "GST1 gene deletion determined by polymerase chain reaction.";
RL   Nucleic Acids Res. 18:3670-3670(1990).
RN   [5]
RP   SEQUENCE OF 1-23.
RX   MEDLINE=86042634; PubMed=3864155;
RA   Mannervik B., Alin P., Guthenberg C., Jensson H., Tahir M.K.,
RA   Warholm M., Joernvall H.;
RT   "Identification of three classes of cytosolic glutathione transferase
RT   common to several mammalian species: correlation between structural
RT   data and enzymatic properties.";
RL   Proc. Natl. Acad. Sci. U.S.A. 82:7202-7206(1985).
RN   [6]
RP   SEQUENCE OF 1-23.
RX   MEDLINE=85154554; PubMed=3979555;
RA   Alin P., Mannervik B., Joernvall H.;
RT   "Structural evidence for three different types of glutathione
RT   transferase in human tissues.";
RL   FEBS Lett. 182:319-322(1985).
RN   [7]
RP   SEQUENCE OF 52-59, AND MASS SPECTROMETRY.
RC   TISSUE=Liver;
RA   Hubbard M.J., McHugh N.J.;
RT   "Human ERp29: Isolation, primary structural characterisation and
RT   two-dimensional gel mapping.";
RL   Electrophoresis 21:3785-3796(2000).
RN   [8]
```

```
RP   X-RAY CRYSTALLOGRAPHY (2.68 ANGSTROMS).
RX   MEDLINE=99130249; PubMed=9930979;
RA   Patskovsky Y.V., Patskovska L.N., Listowsky I.;
RT   "Functions of His107 in the catalytic mechanism of human glutathione
RT   S-transferase hGSTM1a-1a.";
RL   Biochemistry 38:1193-1202(1999).
CC   -!- FUNCTION: CONJUGATION OF REDUCED GLUTATHIONE TO A WIDE NUMBER
CC       OF EXOGENOUS AND ENDOGENOUS HYDROPHOBIC ELECTROPHILES.
CC   -!- CATALYTIC ACTIVITY: RX + GLUTATHIONE = HX + R-S-GLUTATHIONE.
CC   -!- SUBUNIT: HOMODIMER.
CC   -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC   -!- TISSUE SPECIFICITY: THIS IS A LIVER ISOZYME.
CC   -!- POLYMORPHISM: THERE ARE TWO ALLELES; GSTM1A AND GSTM1B WHICH
CC       DIFFER IN POSITION 172.
CC   -!- SIMILARITY: BELONGS TO THE GST SUPERFAMILY. MU FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; X08020; CAA30821.1; -.
DR   EMBL; J03817; AAA59203.1; -.
DR   EMBL; X68676; CAA48636.1; -.
DR   EMBL; X51451; CAA35817.1; -.
DR   PIR; S01719; S01719.
DR   PIR; B22457; B22457.
DR   PIR; E24735; E24735.
DR   PDB; 1GTU; 02-FEB-99.
DR   MIM; 138350; -.
DR   InterPro; IPR000521; -.
DR   InterPro; IPR003081; -.
DR   Pfam; PF00043; GST; 1.
DR   PRINTS; PR01267; GSTRNSFRASEM.
KW   Transferase; Multigene family; Polymorphism; 3D-structure.
FT   INIT_MET      0      0
FT   VARIANT     172    172       K -> N (IN ALLELE B).
FT                                /FTId=VAR_003617.
FT   CONFLICT     43     43       S -> T (IN REF. 3).
SQ   SEQUENCE   217 AA;  25580 MW;  A8BB0E713BA75EAB CRC64;

  Query Match             70.6%;  Score 36;  DB 1;  Length 217;
  Best Local Similarity   85.7%;  Pred. No. 6.4;
  Matches    6;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       3 FPNAPYL 9
           ||| |||
Db      56 FPNLPYL 62


RESULT 10
GTM1_MOUSE
ID   GTM1_MOUSE     STANDARD;      PRT;   217 AA.
AC   P10649;
DT   01-JUL-1989 (Rel. 11, Created)
DT   01-JUL-1989 (Rel. 11, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   GLUTATHIONE S-TRANSFERASE GT8.7 (EC 2.5.1.18) (GST 1-1) (GST CLASS-
DE   MU).
GN   GSTM1.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=88330838; PubMed=3417659;
RA   Pearson W.R., Reinhart J., Sisk S.C., Anderson K.S., Adler P.N.;
RT   "Tissue-specific induction of murine glutathione transferase mRNAs by
RT   butylated hydroxyanisole.";
RL   J. Biol. Chem. 263:13324-13332(1988).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90094327; PubMed=2689439;
RA   Townsend A.J., Goldsmith M.E., Pickett C.B., Cowan K.H.;
RT   "Isolation, characterization, and expression in Escherichia coli of
RT   two murine Mu class glutathione S-transferase cDNAs homologous to the
RT   rat subunits 3 (Yb1) and 4 (Yb2).";
RL   J. Biol. Chem. 264:21582-21590(1989).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93290350; PubMed=8512323;
RA   Reinhart J., Pearson W.R.;
RT   "The structure of two murine class-mu glutathione transferase genes
RT   coordinately induced by butylated hydroxyanisole.";
RL   Arch. Biochem. Biophys. 303:383-393(1993).
RN   [4]
RP   PRELIMINARY SEQUENCE OF 1-40.
RX   MEDLINE=83109018; PubMed=6822548;
RA   Pearson W.R., Windle J.J., Morrow J.F., Benson A.M., Talalay P.;
RT   "Increased synthesis of glutathione S-transferases in response to
RT   anticarcinogenic antioxidants. Cloning and measurement of messenger
RT   RNA.";
RL   J. Biol. Chem. 258:2052-2062(1983).
RN   [5]
RP   PRELIMINARY SEQUENCE OF 1-24.
RX   MEDLINE=86042634; PubMed=3864155;
RA   Mannervik B., Alin P., Guthenberg C., Jensson H., Tahir M.K.,
RA   Warholm M., Joernvall H.;
RT   "Identification of three classes of cytosolic glutathione transferase
RT   common to several mammalian species: correlation between structural
RT   data and enzymatic properties.";
RL   Proc. Natl. Acad. Sci. U.S.A. 82:7202-7206(1985).
RN   [6]
RP   CHARACTERIZATION.
RC   STRAIN=CD-1; TISSUE=Liver;
RX   MEDLINE=96189427; PubMed=8605288;
RA   Mitchell A.E., Morin D., Lame M.W., Jones A.D.;
RT   "Purification, mass spectrometric characterization, and covalent
RT   modification of murine glutathione S-transferases.";
RL   Chem. Res. Toxicol. 8:1054-1062(1995).
CC   -!- FUNCTION: CONJUGATION OF REDUCED GLUTATHIONE TO A WIDE NUMBER
CC       OF EXOGENOUS AND ENDOGENOUS HYDROPHOBIC ELECTROPHILES.
CC   -!- CATALYTIC ACTIVITY: RX + GLUTATHIONE = HX + R-S-GLUTATHIONE.
CC   -!- SUBUNIT: HOMODIMER.
CC   -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC   -!- MASS SPECTROMETRY: MW=25838.4; MW_ERR=2; METHOD=ELECTROSPRAY.
CC   -!- SIMILARITY: BELONGS TO THE GST SUPERFAMILY. MU FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; J03952; AAA37747.1; -.
DR   EMBL; J04632; AAA37705.1; -.
DR   EMBL; L13448; -; NOT_ANNOTATED_CDS.
DR   PIR; A20831; A20831.
DR   PIR; A28946; A28946.
DR   PIR; I24735; I24735.
DR   PIR; A34159; A34159.
DR   PIR; S33860; S33860.
DR   HSSP; P04905; 6GSY.
DR   SWISS-2DPAGE; P10649; MOUSE.
DR   MGD; MGI:95860; Gstm1.
DR   InterPro; IPR000521; -.
DR   InterPro; IPR003081; -.
DR   Pfam; PF00043; GST; 1.
DR   PRINTS; PR01267; GSTRNSFRASEM.
```

```
KW   Transferase; Multigene family.
FT   INIT_MET      0      0
SQ   SEQUENCE   217 AA;  25839 MW;  A1EE3938F590B829 CRC64;


  Query Match             70.6%;  Score 36;  DB 1;  Length 217;
  Best Local Similarity   85.7%;  Pred. No. 6.4;
  Matches    6;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       3 FPNAPYL 9
           ||| |||
Db      56 FPNLPYL 62


RESULT  11
GTM1_RAT
ID   GTM1_RAT       STANDARD;      PRT;   217 AA.
AC   P04905;
DT   13-AUG-1987 (Rel. 05, Created)
DT   01-AUG-1988 (Rel. 08, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   GLUTATHIONE S-TRANSFERASE YB1 (EC 2.5.1.18) (CHAIN 3) (GST M1-1)
DE   (GST CLASS-MU).
GN   GSTM1.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A. (CLONE PGTR200).
RX   MEDLINE=86312882; PubMed=2875437;
RA   Lai H.-C.J., Grove G., Tu C.-P.D.;
RT   "Cloning and sequence analysis of a cDNA for a rat liver glutathione
RT   S-transferase Yb subunit.";
RL   Nucleic Acids Res. 14:6101-6114(1986).
RN   [2]
RP   SEQUENCE FROM N.A. (CLONE PGTA/C44).
RX   MEDLINE=86033768; PubMed=3840477;
RA   Ding G.J.-F., Lu A.Y.H., Pickett C.B.;
RT   "Rat liver glutathione S-transferases. Nucleotide sequence analysis
RT   of a Yb1 cDNA clone and prediction of the complete amino acid
RT   sequence of the Yb1 subunit.";
RL   J. Biol. Chem. 260:13268-13271(1985).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=86224097; PubMed=3011803;
RA   Ding G.J.-F., Ding V.D.-H., Rodkey J.A., Bennett C.D., Lu A.Y.H.,
RA   Pickett C.B.;
RT   "Rat liver glutathione S-transferases. DNA sequence analysis of a Yb2
RT   cDNA clone and regulation of the Yb1 and Yb2 mRNAs by
RT   phenobarbital.";
RL   J. Biol. Chem. 261:7952-7957(1986).
RN   [4]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87308179; PubMed=3040722;
RA   Chang C., Saltzman A.G., Sorensen N.S., Hiipakka R.A., Liao S.;
RT   "Identification of glutathione S-transferase Yb1 mRNA as the
RT   androgen-repressed mRNA by cDNA cloning and sequence analysis.";
RL   J. Biol. Chem. 262:11901-11903(1987).
RN   [5]
RP   SEQUENCE OF 1-23.
RC   STRAIN=WISTAR; TISSUE=Olfactory epithelium;
RX   MEDLINE=93277499; PubMed=8503873;
RA   Ben-Arie N., Khen M., Lancet D.;
RT   "Glutathione S-transferases in rat olfactory epithelium:
RT   purification, molecular properties and odorant biotransformation.";
RL   Biochem. J. 292:379-384(1993).
RN   [6]
RP   MUTAGENESIS OF CYS-86.
RX   MEDLINE=91354218; PubMed=1883338;
RA   Hsieh J.-C., Huang S.-C., Chen W.-L., Lai Y.-C., Tam M.F.;
RT   "Cysteine-86 is not needed for the enzymic activity of glutathione S-
RT   transferase 3-3.";
```

```
RL   Biochem. J. 278:293-297(1991).
RN   [7]
RP   X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).
RX   MEDLINE=93041702; PubMed=1420139;
RA   Ji X., Zhang P., Armstrong R.N., Gilliland G.L.;
RT   "The three-dimensional structure of a glutathione S-transferase from
RT   the mu gene class. Structural analysis of the binary complex of
RT   isoenzyme 3-3 and glutathione at 2.2-A resolution.";
RL   Biochemistry 31:10169-10184(1992).
RN   [8]
RP   X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).
RA   Fu J.-H., Rose J., Tam M.F., Wang B.-C.;
RT   "New crystal forms of a mu-class glutathione S-transferase from rat
RT   liver.";
RL   Acta Crystallogr. D 50:219-224(1994).
RN   [9]
RP   X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).
RX   MEDLINE=94153886; PubMed=8110735;
RA   Ji X., Johnson W.W., Sesay M.A., Dickert L., Prasad S.M., Ammon H.L.,
RA   Armstrong R.N., Gilliland G.L.;
RT   "Structure and function of the xenobiotic substrate binding site of a
RT   glutathione S-transferase as revealed by X-ray crystallographic
RT   analysis of product complexes with the diastereomers of 9-(S-
RT   glutathionyl)-10-hydroxy-9,10-dihydrophenanthrene.";
RL   Biochemistry 33:1043-1052(1994).
CC   -!- FUNCTION: CONJUGATION OF REDUCED GLUTATHIONE TO A WIDE NUMBER
CC       OF EXOGENOUS AND ENDOGENOUS HYDROPHOBIC ELECTROPHILES.
CC   -!- FUNCTION: THE OLFACTORY GST MAY BE CRUCIAL FOR THE ACUITY OF THE
CC       OLFACTORY PROCESS.
CC   -!- CATALYTIC ACTIVITY: RX + GLUTATHIONE = HX + R-S-GLUTATHIONE.
CC   -!- SUBUNIT: HOMODIMER OR HETERODIMER.
CC   -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC   -!- MISCELLANEOUS: RAT LIVER GLUTATHIONE S-TRANSFERASES (GST) CONSIST
CC       OF BINARY COMBINATIONS OF THREE MAJOR CLASSES OF SUBUNITS
CC       DESIGNATED: YA (25 KDA), YB (27 KDA) AND YC (28 KDA).
CC   -!- MISCELLANEOUS: YB SUBCLASS SELECTIVELY BINDS STEROID HORMONES.
CC   -!- SIMILARITY: BELONGS TO THE GST SUPERFAMILY. MU FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; X04229; CAA27811.1; -.
DR   EMBL; M11719; AAA41287.1; -.
DR   EMBL; J02810; AAA41293.1; -.
DR   PIR; A24085; A24085.
DR   PIR; A25510; A25510.
DR   PIR; A29794; A29794.
DR   PIR; S17167; S17167.
DR   PDB; 1GSB; 31-OCT-93.
DR   PDB; 1GSC; 31-OCT-93.
DR   PDB; 2GST; 31-OCT-93.
DR   PDB; 3GST; 31-JAN-94.
DR   PDB; 4GST; 31-OCT-93.
DR   PDB; 5GST; 31-OCT-93.
DR   PDB; 6GST; 08-NOV-96.
DR   PDB; 6GSU; 08-NOV-96.
DR   PDB; 6GSV; 08-NOV-96.
DR   PDB; 6GSW; 08-NOV-96.
DR   PDB; 6GSX; 08-NOV-96.
DR   PDB; 6GSY; 08-NOV-96.
DR   PDB; 5FWG; 27-JAN-99.
DR   InterPro; IPR000521; -.
DR   InterPro; IPR003081; -.
DR   Pfam; PF00043; GST; 1.
DR   PRINTS; PR01267; GSTRNSFRASEM.
KW   Transferase; Multigene family; 3D-structure; Olfaction.
FT   INIT_MET      0      0
FT   MUTAGEN      86     86       C->S: NO CHANGE IN ACTIVITY.
```

```
FT   CONFLICT    168    168       I -> N (IN REF. 3).
FT   CONFLICT    198    199       KS -> NC (IN REF. 2).
FT   STRAND        2      7
FT   TURN         11     13
FT   HELIX        14     22
FT   TURN         23     24
FT   STRAND       27     32
FT   TURN         37     39
FT   HELIX        43     46
FT   TURN         47     50
FT   STRAND       61     64
FT   TURN         65     66
FT   STRAND       67     70
FT   HELIX        72     82
FT   TURN         83     84
FT   HELIX        90    114
FT   TURN        115    115
FT   TURN        117    118
FT   HELIX       119    128
FT   TURN        129    129
FT   HELIX       130    141
FT   TURN        142    143
FT   TURN        154    154
FT   HELIX       155    169
FT   TURN        171    176
FT   HELIX       178    188
FT   TURN        189    189
FT   HELIX       191    196
FT   TURN        197    198
FT   TURN        200    201
FT   TURN        210    211
SQ   SEQUENCE   217 AA;  25782 MW;  2ACE8D49DA785118 CRC64;

  Query Match             70.6%;  Score 36;  DB 1;  Length 217;
  Best Local Similarity   85.7%;  Pred. No. 6.4;
  Matches    6;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        3 FPNAPYL 9
            ||| |||
Db       56 FPNLPYL 62

RESULT 12
GTM2_HUMAN
ID   GTM2_HUMAN     STANDARD;      PRT;   217 AA.
AC   P28161;
DT   01-APR-1993 (Rel. 25, Created)
DT   01-APR-1993 (Rel. 25, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   GLUTATHIONE S-TRANSFERASE MU 2 (EC 2.5.1.18) (GSTM2-2) (GST CLASS-MU).
GN   GSTM2 OR GST4.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91239584; PubMed=2034681;
RA   Vorachek W.R., Pearson W.R., Rule G.S.;
RT   "Cloning, expression, and characterization of a class-mu glutathione
RT   transferase from human muscle, the product of the GST4 locus.";
RL   Proc. Natl. Acad. Sci. U.S.A. 88:4443-4447(1991).
RN   [2]
RP   X-RAY CRYSTALLOGRAPHY (1.85 ANGSTROMS).
RX   MEDLINE=94238693; PubMed=8182750;
RA   Raghunathan S., Chandross R.J., Kretsinger R.H., Allison T.J.,
RA   Penington C.J., Rule G.S.;
RT   "Crystal structure of human class mu glutathione transferase GSTM2-2.
RT   Effects of lattice packing on conformational heterogeneity.";
RL   J. Mol. Biol. 238:815-832(1994).
CC   -!- FUNCTION: CONJUGATION OF REDUCED GLUTATHIONE TO A WIDE NUMBER
CC       OF EXOGENOUS AND ENDOGENOUS HYDROPHOBIC ELECTROPHILES.
CC   -!- CATALYTIC ACTIVITY: RX + GLUTATHIONE = HX + R-S-GLUTATHIONE.
CC   -!- SUBUNIT: HOMODIMER.
CC   -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC   -!- TISSUE SPECIFICITY: THIS IS A MUSCLE ISOZYME.
CC   -!- SIMILARITY: BELONGS TO THE GST SUPERFAMILY. MU FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; M63509; AAA60963.1; -.
DR   PIR; A39375; A39375.
DR   PDB; 1HNA; 31-JAN-94.
DR   PDB; 1HNB; 31-JAN-94.
DR   PDB; 1HNC; 31-JAN-94.
DR   PDB; 2GTU; 02-MAR-99.
DR   MIM; 138380; -.
DR   InterPro; IPR000521; -.
DR   InterPro; IPR003081; -.
DR   Pfam; PF00043; GST; 1.
DR   PRINTS; PR01267; GSTRNSFRASEM.
KW   Transferase; Multigene family; 3D-structure.
FT   INIT_MET      0      0
FT   STRAND        2      5
FT   STRAND        7      7
FT   TURN         11     12
FT   HELIX        13     22
FT   TURN         23     24
FT   STRAND       27     29
FT   STRAND       32     32
FT   STRAND       35     35
FT   TURN         38     39
FT   STRAND       41     41
FT   HELIX        43     49
FT   TURN         50     52
FT   STRAND       61     64
FT   TURN         65     66
FT   STRAND       67     70
FT   HELIX        72     81
FT   TURN         82     84
FT   HELIX        90    113
FT   TURN        114    115
FT   TURN        117    118
FT   HELIX       119    127
FT   TURN        128    129
FT   HELIX       130    141
FT   TURN        142    143
FT   STRAND      146    146
FT   TURN        147    147
FT   STRAND      148    149
FT   TURN        150    150
FT   STRAND      151    151
FT   HELIX       154    169
FT   HELIX       171    176
FT   HELIX       178    189
FT   HELIX       191    197
FT   TURN        198    198
FT   TURN        214    215
SQ   SEQUENCE   217 AA;  25613 MW;  25603A909482CA39 CRC64;

  Query Match             70.6%;  Score 36;  DB 1;  Length 217;
  Best Local Similarity   85.7%;  Pred. No. 6.4;
  Matches    6;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        3 FPNAPYL 9
            ||| |||
Db       56 FPNLPYL 62
```

```
RESULT  13
GTM2_MOUSE
ID   GTM2_MOUSE     STANDARD;      PRT;   217 AA.
AC   P15626;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   GLUTATHIONE S-TRANSFERASE 5 (EC 2.5.1.18) (GST 5-5) (GST CLASS-MU).
GN   GSTM2.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90094327; PubMed=2689439;
RA   Townsend A.J., Goldsmith M.E., Pickett C.B., Cowna K.H.;
RT   "Isolation, characterization, and expression in Escherichia coli of
RT   two murine Mu class glutathione S-transferase cDNAs homologous to the
RT   rat subunits 3 (Yb1) and 4 (Yb2).";
RL   J. Biol. Chem. 264:21582-21590(1989).
CC   -!- FUNCTION: CONJUGATION OF REDUCED GLUTATHIONE TO A WIDE NUMBER
CC       OF EXOGENOUS AND ENDOGENOUS HYDROPHOBIC ELECTROPHILES.
CC   -!- CATALYTIC ACTIVITY: RX + GLUTATHIONE = HX + R-S-GLUTATHIONE.
CC   -!- SUBUNIT: HOMODIMER.
CC   -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC   -!- SIMILARITY: BELONGS TO THE GST SUPERFAMILY. MU FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; J04696; AAA37706.1; -.
DR   PIR; B34159; B34159.
DR   HSSP; P04905; 6GST.
DR   MGD; MGI:95861; Gstm2.
DR   InterPro; IPR000521; -.
DR   InterPro; IPR003081; -.
DR   Pfam; PF00043; GST; 1.
DR   PRINTS; PR01267; GSTRNSFRASEM.
KW   Transferase; Multigene family.
FT   INIT_MET      0      0
SQ   SEQUENCE   217 AA;  25585 MW;  A3979D4435EF08AE CRC64;

  Query Match             70.6%;  Score 36;  DB 1;  Length 217;
  Best Local Similarity   85.7%;  Pred. No. 6.4;
  Matches    6;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       3 FPNAPYL 9
           ||| |||
Db      56 FPNLPYL 62


RESULT  14
GTM2_RAT
ID   GTM2_RAT       STANDARD;      PRT;   217 AA.
AC   P08010;
DT   01-AUG-1988 (Rel. 08, Created)
DT   01-AUG-1988 (Rel. 08, Last sequence update)
DT   01-NOV-1995 (Rel. 32, Last annotation update)
DE   GLUTATHIONE S-TRANSFERASE YB2 (EC 2.5.1.18) (CHAIN 4) (GST CLASS-MU).
GN   GSTM2.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
```

```
RP   SEQUENCE FROM N.A.
RX   MEDLINE=88298790; PubMed=3403534;
RA   Lai H.-C.J., Qian B., Grove G., Tu C.-P.D.;
RT   "Gene expression of rat glutathione S-transferases. Evidence for gene
RT   conversion in the evolution of the Yb multigene family.";
RL   J. Biol. Chem. 263:11389-11395(1988).
RN   [2]
RP   SEQUENCE.
RC   STRAIN=SPRAGUE-DAWLEY;
RX   MEDLINE=86192461; PubMed=3699019;
RA   Alin P., Mannervik B., Joernvall H.;
RT   "Cytosolic rat liver glutathione transferase 4-4. Primary structure
RT   of the protein reveals extensive differences between homologous
RT   glutathione transferases of classes alpha and mu.";
RL   Eur. J. Biochem. 156:343-350(1986).
RN   [3]
RP   SEQUENCE OF 24-217 FROM N.A.
RX   MEDLINE=86224097; PubMed=3011803;
RA   Ding G.J.-F., Ding V.D.-H., Rodkey J.A., Bennett C.D., Lu A.Y.H.,
RA   Pickett C.B.;
RT   "Rat liver glutathione S-transferases. DNA sequence analysis of a Yb2
RT   cDNA clone and regulation of the Yb1 and Yb2 mRNAs by
RT   phenobarbital.";
RL   J. Biol. Chem. 261:7952-7957(1986).
RN   [4]
RP   SEQUENCE OF 32-217 FROM N.A.
RX   MEDLINE=87008619; PubMed=3020050;
RA   Lai H.-C.J., Tu C.-P.D.;
RT   "Rat glutathione S-transferases supergene family. Characterization of
RT   an anionic Yb subunit cDNA clone.";
RL   J. Biol. Chem. 261:13793-13799(1986).
RN   [5]
RP   SEQUENCE OF 1-21.
RC   STRAIN=WISTAR; TISSUE=Olfactory epithelium;
RX   MEDLINE=93277499; PubMed=8503873;
RA   Ben-Arie N., Khen M., Lancet D.;
RT   "Glutathione S-transferases in rat olfactory epithelium:
RT   purification, molecular properties and odorant biotransformation.";
RL   Biochem. J. 292:379-384(1993).
CC   -!- FUNCTION: CONJUGATION OF REDUCED GLUTATHIONE TO A WIDE NUMBER
CC       OF EXOGENOUS AND ENDOGENOUS HYDROPHOBIC ELECTROPHILES.
CC   -!- FUNCTION: THE OLFACTORY GST MAY BE CRUCIAL FOR THE ACUITY OF THE
CC       OLFACTORY PROCESS.
CC   -!- CATALYTIC ACTIVITY: RX + GLUTATHIONE = HX + R-S-GLUTATHIONE.
CC   -!- SUBUNIT: HOMODIMER OR HETERODIMER.
CC   -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC   -!- MISCELLANEOUS: RAT LIVER GLUTATHIONE S-TRANSFERASES (GST) CONSIST
CC       OF BINARY COMBINATIONS OF THREE MAJOR CLASSES OF SUBUNITS
CC       DESIGNATED: YA (25 KDA), YB (27 KDA) AND YC (28 KDA).
CC   -!- MISCELLANEOUS: YB SUBCLASS SELECTIVELY BINDS STEROID HORMONES.
CC   -!- SIMILARITY: BELONGS TO THE GST SUPERFAMILY. MU FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; J02592; AAA41285.1; -.
DR   EMBL; M13590; AAA42351.1; -.
DR   EMBL; J03914; AAA41296.1; -.
DR   PIR; A25386; XURTG4.
DR   PIR; A26307; A26307.
DR   PIR; B26187; B26187.
DR   HSSP; P04905; 6GSU.
DR   InterPro; IPR000521; -.
DR   InterPro; IPR003081; -.
DR   Pfam; PF00043; GST; 1.
DR   PRINTS; PR01267; GSTRNSFRASEM.
KW   Transferase; Multigene family; Olfaction.
FT   INIT_MET      0      0
```

```
FT   CONFLICT    146    146       W -> S (IN REF. 2).
SQ   SEQUENCE   217 AA;  25571 MW;  F27B3D5831FF789F CRC64;


  Query Match             70.6%;  Score 36;  DB 1;  Length 217;
  Best Local Similarity   85.7%;  Pred. No. 6.4;
  Matches    6;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       3 FPNAPYL 9
           ||| |||
Db      56 FPNLPYL 62


RESULT  15
GTM3_MOUSE
ID   GTM3_MOUSE     STANDARD;      PRT;   217 AA.
AC   P19639;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   GLUTATHIONE S-TRANSFERASE GT9.3 (EC 2.5.1.18) (GST CLASS-MU).
GN   GSTM3.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=88330838; PubMed=3417659;
RA   Pearson W.R., Reinhart J., Sisk S.C., Anderson K.S., Adler P.N.;
RT   "Tissue-specific induction of murine glutathione transferase mRNAs by
RT   butylated hydroxyanisole.";
RL   J. Biol. Chem. 263:13324-13332(1988).
RN   [2]
RP   SEQUENCE OF 1-40 FROM N.A.
RX   MEDLINE=83109018; PubMed=6822548;
RA   Pearson W.R., Windle J.J., Morrow J.F., Benson A.M., Talalay P.;
RT   "Increased synthesis of glutathione S-transferases in response to
RT   anticarcinogenic antioxidants. Cloning and measurement of messenger
RT   RNA.";
RL   J. Biol. Chem. 258:2052-2062(1983).
RN   [3]
RP   SEQUENCE OF 1-25.
RX   MEDLINE=86042634; PubMed=3864155;
RA   Mannervik B., Alin P., Guthenberg C., Jensson H., Tahir M.K.,
RA   Warholm M., Joernvall H.;
RT   "Identification of three classes of cytosolic glutathione transferase
RT   common to several mammalian species: correlation between structural
RT   data and enzymatic properties.";
RL   Proc. Natl. Acad. Sci. U.S.A. 82:7202-7206(1985).
CC   -!- FUNCTION: CONJUGATION OF REDUCED GLUTATHIONE TO A WIDE NUMBER
CC       OF EXOGENOUS AND ENDOGENOUS HYDROPHOBIC ELECTROPHILES.
CC   -!- CATALYTIC ACTIVITY: RX + GLUTATHIONE = HX + R-S-GLUTATHIONE.
CC   -!- SUBUNIT: HOMODIMER.
CC   -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC   -!- SIMILARITY: BELONGS TO THE GST SUPERFAMILY. MU FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; J03953; AAA37748.1; ALT_INIT.
DR   PIR; B20831; B20831.
DR   PIR; B28946; B28946.
DR   PIR; E37520; E37520.
DR   HSSP; P04905; 6GST.
DR   MGD; MGI:106026; Gstm3.
DR   InterPro; IPR000521; -.
DR   InterPro; IPR003081; -.
DR   Pfam; PF00043; GST; 1.
DR   PRINTS; PR01267; GSTRNSFRASEM.
KW   Transferase; Multigene family.
FT   INIT_MET      0      0
SQ   SEQUENCE   217 AA;  25570 MW;  C4D8950EFEA585D8 CRC64;

  Query Match             70.6%;  Score 36;  DB 1;  Length 217;
  Best Local Similarity   85.7%;  Pred. No. 6.4;
  Matches    6;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       3 FPNAPYL 9
           ||| |||
Db      56 FPNLPYL 62


Search completed: October 23, 2001, 13:28:18
Job time: 401 sec
```